

Volume 9, Number 1
2016

ISSN 1807-9792

abstracta

Linguagem, Mente & Ação

<http://abstracta.oa.hhu.de>

In Carnap's Defense: A survey on the concept of a linguistic
framework in Carnap's philosophy
Parzhad Torfehnezhad

Justifying Practical Reasons
Georg Spielthener

O caráter definicional *sui generis* dos predicados tarskianos de verdade
Luciano Vicente

What is Logical about the Logical Interpretation of Probability?
Parzhad Torfehnezhad

Self-Concernment without Self-Reference
Roberto Horácio de Sá Pereira

d|u|p

abstracta

Linguagem, Mente & Ação

ISSN 1807-9792

Volume 9, Number 1
2016

Editors

André Joffily Abath
Leonardo Ribeiro
Gottfried Vosgerau

Executive Editors

Nicolas Lindner
Alex Tillas

Associate Editors

Giuliano Torrenço
José Edgar González Varela

Contents

In Carnap's Defense: A survey on the concept of a linguistic framework in Carnap's philosophy	3
<i>Parzhad Torfehnezhad</i>	
Justifying Practical Reasons	31
<i>Georg Spielthenner</i>	
O caráter definicional <i>sui generis</i> dos predicados tarskianos de verdade	45
<i>Luciano Vicente</i>	
What is Logical about the Logical Interpretation of Probability?	51
<i>Parzhad Torfehnezhad</i>	
Self-Concernment without Self-Reference	69
<i>Roberto Horácio de Sá Pereira</i>	

In Carnap's Defense: A survey on the concept of a linguistic framework in Carnap's philosophy

Parzhad Torfehnezhad

Département de Philosophie
Université de Montréal
Pavillon 2910, boul. Édouard-Montpetit
Montréal, QC, H3C 3J7, Canada
parzhad.torfeh-nezhad@umontreal.ca

Abstract

The main task in this paper is to detail and investigate Carnap's conception of a "linguistic framework" (LF). On this basis, we will see whether Carnap's dichotomies, such as the analytic-synthetic distinction, are to be construed as absolute/fundamental dichotomies or merely as relative dichotomies. I argue for a novel interpretation of Carnap's conception of a LF and, on that basis, will show that, according to Carnap, all the dichotomies to be discussed are relative dichotomies; they depend on conventional decisions concerning the logical syntax of LF. Thus, all of the dichotomies directly hinge on the conception of the LF. The LF's logical structure, in turn, is an immediate consequence of adopting the linguistic doctrine of logical truths. As we will see, no appeal to any of these distinctions is necessary in establishing a LF and all of its components. I will also draw attention to the differences between what Carnap labels a "way of speaking", "language", and "artificial language". Consequently, I will briefly conclude that none of Quine's major objections address the main points of Carnap's theory.

1 Introduction

The Quine-Carnap debate is one of the most well-known debates in the history of modern philosophy. For Carnap, ontological questions like "Is/are there so and so?" are meaningless if they are questions external to what he calls a "linguistic framework" (henceforth LF). In other words, Carnap believes meaningful ontological questions, in general, can only be asked from inside an adopted LF (Carnap, 1950: 914–915). Quine, on the other hand, does not distinguish between external and internal questions. He argues that this distinction is only based on the old, fundamental (absolute) analytic-synthetic distinction. The latter, on Quine's view, is both wrong and useless (Quine, 1951: 43).

In the literature, numerous scholars and commentators devoted at least a portion of their work to reflect on, or discuss, the Quine-Carnap dispute. Some, from among those who defended Carnap's position, think that it is immune from Quine's criticism if one properly analyses the relationships between Carnap's dichotomies. By appealing to what he calls a "metaphorical-literal distinction", Yablo, for example, claims that the association of the internal-external distinction with analytic-synthetic distinction (henceforth ASD) can be freed and shown to be a "non-committal figurative speech", so that even Quine cannot argue against it (Yablo, 1998: 232–233). Similarly, Bird thinks Quine misses Carnap's central points by failing to appreciate the "four-folded" distinction underlying Carnap's internal-external distinction (Bird, 1995).

Some philosophers argue that Quine is equally guilty of the same charge of which he accuses Carnap. Berge (Berge, 1995) rejects Quine's criticism on the basis that his view on reference is quite similar to that of Carnap's. In addition, Burgess (Burgess, 2004) believes that although Quine is right to argue that the internal-external distinction is based on the ASD, he inevitably needs something similar to explain the obviousness of elementary mathematics. Similarly, Koellner (Koellner, preprint), in the context of mathematical truth, fairly defends and justifies Carnap's view on analyticity and mathematical pluralism.

Some do not question the fundamentality of the ASD, but they take the distinction to be representative of other, deeper distinctions. O'Grady (O'Grady, 1999) shows how some scholars mistakenly evaluated Carnap's position. Nevertheless, he appreciates that the dispute could be understood as a dispute about deeper philosophical methods. Also, Lavers (Lavers, 2012) argues that the Quine-Carnap dispute on analyticity stems from their different views on what constitutes a successful explication.

Furthermore, there are philosophers who argue that Carnap's neutral ontological position is achievable via some modifications. Friedman (Friedman, 2009), for example, argues that if Carnap's scientific theory is understood in conjunction with Ramsey's sentences, the neutrality of Carnap's ontological position can be restored. Others, like Grice and Strawson (Grice, 1956), argue that Quine's criticism is simply not sufficient to reject the ASD.

Some philosophers take a different approach altogether. They tend to evaluate Carnap's general philosophy of science without engaging the debate about the dichotomies directly. Interpreting Carnap as an instrumentalist, Howard Stein (Stein, 1989) discusses the legitimacy, importance, and productivity of both realism and instrumentalism from the perspective of history and philosophy of science. He then evaluates the debate as a productive example of a realist-instrumentalist debate. On the other hand, Hintikka (Hintikka, 1992) believes that focusing on Carnap's dichotomies would not be helpful in illuminating the real problem about the general dynamics of Carnap's thoughts. Hintikka argued that although it may appear to be the case that Carnap takes language as a calculus, he actually maintains the idea of the universality of language, and does not regard language as a mere calculus. This, for Hintikka, is the main problem.¹

As we have seen in all of the above-mentioned examples, one of the questions that has not received enough attention is the question of whether these dichotomies are treated as fundamental (absolute) or relative. Is the ASD or internal-external distinction, regardless of their relation to each other, understood by Carnap to be *fundamental/absolute* distinctions? Or are they treated by Carnap as *relative* distinctions that are decidable only after we adopt a particular LF? In other words, is it akin to the relative distinction between east and west after we agree on the particular geographic region in question, or to the absolute distinction between "to be" and "not to be"? One of the main questions in this paper concerns whether these dichotomies are understood absolutely, i. e., whether they are treated as fundamental dichotomies. To this question, I will answer in the negative. Consistent with the given interpretation of LF in this paper, I will argue that, according to Carnap, all the mentioned dichotomies are *relative* dichotomies. They turn on our conventional decision concerning the logical syntax of the LF. The conception of all dichotomies directly hinge upon the conception of the LF, and the LF's

¹The mentioned references are only few examples of a large literature on this topic. Although listing and discussing all of them is neither possible nor intended in this paper, here are some other examples that one may want to consult: (Price, 2009), (Price, 1997), (Psillos, 2000), (Hillier, 2009), (George, 2000), (Akiba, 1995), (Sober & Hylton, 2000), (McDermott, 2001), (Hempel, 1973), (Psillos, 2000), (Oberdan, 1992), (Haack, 1993), (Tsou, 2003), (Peacock, 2011), (Arnold & Shapiro, 2007), (Soames, 2009), (Friedman, 2000), (Friedman & Creath, 2007), (Friedman, 1999), (A. W. Richardson, 2003), (A. Richardson, 2007), (Awodey, 2007).

logical structure, in turn, is an immediate consequence of adopting the linguistic doctrine of logical truths.

As we will see, no appeal to any of these distinctions is necessary to establish a LF and all of its components. In my view, all the distinctions become immediate simply by accepting that there is such a thing as a LF as described by Carnap. The term "LF" has been used by Rudolf Carnap in his famous paper "Empiricism, Semantics and Ontology" (Carnap, 1950) (henceforth ESO), and he has rarely used it elsewhere. This fact, in my view, caused many misinterpretations of Carnap's and his critics' understanding of the term. Though I will highlight some of these seemingly wrong interpretations, the goal of this paper is not to criticize them. Instead, I will present the interpretation of the term "LF" by scrutinizing some of Carnap's works other than ESO and especially his contribution to the "Encyclopaedia of Unified Sciences" in 1939 (Carnap, 1939). This may lead to a consensus on what Carnap means by a LF.

In short, I will generally characterize Carnap's conception of a LF as a factual-conventional hierarchy of assertions (or strings of signs) that is subjected to certain rules for delivering meaning. The rules could primarily be constructed (or recognized) from purely factual statements up to the purely conventional statements, and could equally be constructed the other way around, i. e., from purely conventional statements of a calculus down to purely factual statements of a newly interpreted language.

In section two, I will describe the grounds upon which language became a central point in Carnap's philosophy. Following this, I will briefly discuss the development of his conception of language, from his view in *Aufbau* (Carnap, 1967)² to his view in ESO (Carnap, 1950). Of course, in this section, my only concern will be Carnap's conception of language relative to his position on logic. This section will help us have a better idea of the basis upon which the "Linguistic Doctrine of Logical Truths" (i. e., logical truths are true by linguistic convention; henceforth LD) was adopted. As we will see, in Carnap's former view, logic is regarded as a representative system directly attached to our explanations of the world. According to this view one may conclude that the world, as we explain it, should have an underlying logical structure. But in Carnap's later view, he notices and legitimizes some sort of invention (in the form of conventions) in the middle of the former process of investigating (or constructing) logical forms of factual statements. By adding the conventionality factor to his theory, Carnap diverges from Wittgenstein. Therefore, the supposed "logical structure of the world" could no longer be the mirror image of the structure of the world. Later on, we will see how Carnap thinks this slight modification makes room for the equally legitimate concept of what is now known as an artificial language (as opposed to a natural language).

In the next section, the hierarchy of abstractions will be presented. There, I will clarify Carnap's later position about language by summarizing Carnap's "Foundations of Logic and Mathematics" (Carnap, 1939) and his views on the way to construct a language system and perform a linguistic analysis. In this section, I will present what Carnap calls a "language system" which, in my view, essentially bears no difference with what he later calls a LF in ESO.

In section four, I will give a detailed explanation of Carnap's two methods for constructing language systems. We will also see the basis of two possible and yet different changes in a

²Originally published as "*Der logische Aufbau der Welt*", Berlin, Benary, 1928. Although conventionalism is present in the *Aufbau*, some scholars believe Carnap is not explicit about this concept there (Runggaldier, 1984: 11).

framework parallel to the applications of these methods: the one that does not alter the logical fabric of the framework, and the one that does.

In section five, we will see how changes could basically be introduced at different levels of abstraction in order to produce moderately or radically different frameworks. In this section we will see what it means to have an “artificial language” as opposed to having different “ways of speaking”.

In section six, I will turn to Quine’s objection after clarifying Carnap’s conception of analyticity in light of what I argue in the previous sections. According to the given explanations, one may realize how Quine’s major objections miss the main points of Carnap’s theory. I will also argue that, as far as it concerns Carnap’s first method, Quine and Carnap are in complete agreement. The disagreement appears only where Carnap considers his second method to be as legitimate as his first. Quine, on the other hand, completely rejects this idea. He argues that the difference between natural and artificial languages (as well as the difference between external and internal questions to a LF) can only be established upon the acceptance of the useless ASD. I will argue that both distinctions directly hinge upon the conception of a LF, which in turn, is immediate by accepting LD.

I will then conclude that Carnap’s conception of a LF is immediate and unobjectionable following the admission of LD. Moreover, Carnap’s distinctions cannot be construed as absolute distinctions. I also show that Carnap’s model for language analysis is more fruitful and constructive compared to Quine’s. The latter, on my view, is more in accordance with traditional ways of thinking about philosophical problems.

2 Historical Background

In this section I will gloss over some historical background in order to elucidate why the notion of language is such a central point in Carnap’s philosophy and why LD becomes such an important doctrine among the neo-empiricists of the Vienna Circle.

In the following quotes, Carnap speaks about his general view on the world-language relationship and his view on the specific position of logic with regard to language. He speaks of both in connection with the ideas of two important figures, Wittgenstein and Neurath:³

For me personally, Wittgenstein was perhaps the philosopher who, besides Russell and Frege, had the greatest influence on my thinking. The most important insight I gained from his work was the conception that the truth of logical statements is based only on their logical structure and on the meaning of the terms. (Schilpp, 1963: 24).

We [in Vienna Circle] read in Wittgenstein’s book that certain things show themselves but cannot be said; for example the logical structure of sentences and the relation between the language and the world. In opposition to this view, first tentatively, then more and more clearly, our conception developed that it is possible to talk meaningfully about language and about the relation between a sentence and the fact described. Neurath emphasized these facts in order to reject the view that there is something “higher”, something mysterious, “spiritual”, in language, a view which was prominent in German philosophy. I agreed with him, but pointed out that only the structural pattern,

³It is well-known among Carnap scholars that Carnap’s thoughts, in general, were influenced by many figures such as Frege, Hilbert, Russell, Tarski, Gödel, and others. Yet, the ideas of Wittgenstein and Neurath were more directly concerned with the concept of language than Carnap’s more significant influences.

not the physical properties of the ink marks, were relevant for the function of language. (Schilpp, 1963: 28).

According to Carnap (Schilpp, 1963: 52) it is not possible for Wittgenstein to talk about language in isolation. It is also apparent from the last couple of verses of *Tractatus* that speaking of propositions and rules of language in total separation from where they are being employed is meaningless. Wittgenstein is clear that language appears to be the unique and correlated picture of the world (Wittgenstein, 1958: §95), and that he considers logic as the underlying and hidden "essence of language" (Ibid, §97). For Wittgenstein, in short, language is a tool for revealing some structure *of the world* via representation.

This instrumental role of language, which brings about logic as a representative system, seems to regard language with a different ontological status than that of the rest of the actual world. This seems to be the problem with this view. On this view, language is something by which we, for instance, explain the world. Language is one thing and the world is another. Language is a tool we use to satisfy a purpose. The question, then, is whether or not the two are ontologically distinct. The problem gets worse when we start thinking *about* logic. On the one hand, we start off our search for logic and get to the "essence of language" from accidental linguistic statements. Therefore, we have to acknowledge some sort of dependency between logic and language. On the other hand, we have to say logic or, as Wittgenstein put it, "the rules of possibilities", is totally independent of all language forms. Accordingly, one has to accept a very mysterious status for logic and language with respect to the rest of the world.

Carnap departs from Wittgenstein at exactly this point; unlike Wittgenstein, talking about language in isolation is possible for Carnap because language itself is a worldly object. In agreement with Neurath, along with other members of Vienna Circle, Carnap admits the possibility of speaking *about* language in isolation (Schilpp, 1963: 52). Unlike Wittgenstein, Neurath considers language as something within the world, not something that refers to the world from the outside (Schilpp, 1963: 28). This view of language is one of the most important turns in Carnap's philosophy (Ibid). Language can still preserve its instrumental role, but now it is a tool that works within a system and not outside of it. To give an analogy, although we may deem red blood cells as instruments or tools for transporting oxygen across the body, they are still parts of the human body. The case is different when we consider instruments for constructing buildings, for example. They are tools that are no longer part of the building after its construction. Tools, in this latter sense, have an ontological status over and above the building (just like language and logic in Wittgenstein's view, which have a distinct status over and above the world). In the former case, red blood cells do not bear such a status. Similarly, we may still consider language as an instrument to talk about the world, but, at the same time, language itself is an object of the world that bears a special relationship to other objects.

According to Carnap (Schilpp, 1963: 28), it was this idea that led him to consider what he later called the "logical syntax of language". Centrality of language also helped Carnap take more radical positions against traditional metaphysics, and adopt a more neutral attitude toward "the various philosophical forms of language", e. g., realism, idealism and the like (Schilpp, 1963:17–18, 24). Carnap formulated this neutral attitude in the form of a "principle of tolerance" in his "Logical Syntax of Language" (Carnap, 1937). Now, in settling the mentioned philosophical controversies such as the realist-nominalist debate (which was caused by the diverse use of language), our concerns are to first look at the syntactical properties of the various forms of language, and secondly, the "practical reasons for preferring one or the other form for given purposes" (Schilpp, 1963: 54). Construing philosophical problems as metalinguistic

problems as opposed to linguistic ones is obvious when Carnap explains his major motivation for adopting the syntactic method:

In our discussions in the Vienna Circle it had turned out that any attempt at formulating more precisely the philosophical problems in which we were interested ended up with problems of the logical analysis of language. Since in our view the issue in philosophical problems concerned the language, not the world, these problems should be formulated, not in the object language, but in the meta-language. (Schilpp, 1963: 54)

It might be fair to say that the idea of considering language as an object within the world and, hence, the possibility of talking about language in isolation, were the main motives in formulating LD: logical truths are true by linguistic convention. The adoption of this doctrine was, of course, an established point of consensus among Carnap and other members of Vienna Circle, although Carnap was not completely in agreement with this formulation of the doctrine (Schilpp, 1963: 914). The acceptance of the doctrine immediately implies a linguistic-based and conventional⁴ nature of the logical structure (a LF) that can be revealed via a complete analysis of language. Any theory that provides descriptions of the steps involved in completing such an analysis, as well as explaining all properties, features, and rules involved in taking these steps eventually (and inevitably), proposes or describes the characterizations of a framework according to which one makes assertions. Carnap's attempt to propose such a theory is the subject matter of the following section.

A philosophical linguistic analysis, in general, is concerned with methods of clarifying concepts behind the terms of the ordinary language with respect to the structures in which the terms are being used; one may simply call the methods of this sort an "explication". The notion of a linguistic framework, evidently, is not only of great importance in his linguistic analysis, but is also directly related to the subject matter of Carnap's overall philosophy.

3 Linguistic Framework and its Components

So far, we may summarize the implications of adopting LD as follows:

1. Language has a (logical) structure.
2. In the very *first* attempt of investigating such a structure there has to be a language in place (as an object).
3. Making conventions is part and parcel of such an investigation.

The main question now is how we can investigate the mentioned structure of the language. How does logic (logical structures) emerge? How is it differentiated from the rest of ordinary language? Carnap provides us with a detailed answer (Carnap, 1939), which I will summarize in this section. For Carnap, language is inclusive of a vast array of "communicative signs" (Carnap, 1994: 291–294). The major purpose of Carnap's project, from now on, is to show the ways in which a so-called "scientific language" differs from our ordinary use of language.

⁴Carnap himself would rather not use the term "convention" or "conventional" for fear of giving the impression that there is too much liberty and arbitrariness involved in the process of identifying logical truths. Since this concept becomes clearer in the following section, I use the term as-is and skip the controversy about "convention" or "conventional". In Carnap's own words:

Among the various formulations [...] there are some which today I would no longer regard as psychologically helpful and would therefore avoid. One of them is the characterization of logical truth as based on "linguistic fiat" or "linguistic conventions". [...] The term "linguistic convention" is usually understood in the sense of a more or less arbitrary decision concerning language, such as the choice of either centimeter or inch as a unit of length. (Schilpp, 1963: 914–915)

To put it differently: by what mechanism does a system of scientific statements (in general, science) start to emerge from the context of ordinary statements? It was the work of people like W. C. Morris (e.g., "Foundations of the Theory of Signs") that helped Carnap develop a complete theory of language (Carnap, 1994: 291–294), so that it is inclusive of the entire spectrum of human assertions. The latter ranges from the assertions in ordinary discourses to mathematical and logical assertions. Carnap considers language systems as hierarchical systems consisting of three parts; respectively, from the bottom to the top, these parts are pragmatics, semantics, and syntax⁵. He frequently refers to these three parts in nearly all of his works after 1939 (Carnap, 1939, 1942, 1959, 1994).

Therefore, an analysis of theoretical procedures in science must concern itself with language and its applications. [...] we shall outline an analysis of language and explain the chief factors involved. Three points of view will be distinguished, and accordingly three disciplines applying them, called pragmatics, semantics, and syntax. [...] The complete theory of language has to study all these three components. (Carnap, 1939: 3–4)

These three components have different focuses of attention yet interrelated and, consequently, they lead to different types of research or activity. In pragmatics, the focus is on the world-speaker (world-language) relation. In semantics, what is under investigation is the relation of designation regardless of (or given) the world-language relation (where, for example, we may expand or limit the meaning of a term or phrase in our use of language). Syntax is where we begin to investigate the (logical) structure of language regardless of (or given) the designation relation in semantics. Given that science has its roots in experiencing the actual world, one should keep in mind that the world under investigation in pragmatics is strictly the actual world (see below). Therefore, it consists of a finite amount of objects. One other important point in the subsequent sections, which deals with the methods of constructing a framework for language, is that the language in question is considered to be an instance of actual historical natural languages. Later on, when we talk about the second method of construction, we will consider this topic in light of artificial languages as well.

3.1 Pragmatics

In pragmatics, speakers of the language generate signs for objects, events, relations, properties, etc., in order to communicate inside the language community, understand/explain actual events, construct theories about the world, etc. Carnap considers problems of a factual and empirical nature, which deal with gaining and communicating knowledge, as problems that belong to pragmatics (Carnap, 1942: 250). These problems have to do with the speaker's activities of perception, observation, comparison, registration, confirmation, etc., as far as they lead to (or refer to) knowledge formulated in a language (Carnap, 1942: 245). Pragmatics is where we study methods of testing hypotheses and theories by deriving predictions from them in the form of "observation sentences", and then comparing these predicted results with new

⁵I should note here that the terms "pragmatics", "semantics", and "syntax" have been originally borrowed from the terminology of linguistics, but for Carnap the scope of these terms is broader; they don't have the exact same referent as they do in linguistics (linguistics, as a pragmatic scientific discipline). One should not confuse the scientific, or rather, the more pragmatic uses of the terms, which bear a descriptive nature, with the more theoretical applications of them. The latter is the manner in which Carnap intend to employ them. Linguistics, at the scientific level, is concerned with the study of actualities about an actual language like English. Therefore, semantics and syntax are to be considered as descriptive semantics and descriptive syntax. They eventually yield an English dictionary or an English grammar. At this level of abstraction, we are still at the level of pragmatics (in the Carnapian use of this term). That is why Carnap sometimes has to emphasize the distinction by using expressions such as "logical syntax" or "pure semantics" as opposed to "descriptive syntax" and "descriptive semantics" in order to avoid the confusion. (see Carnap, 1942: 240)

observation sentences: “The outcome of such a procedure of testing a hypothesis is either a confirmation or an infirmation of that hypothesis, or, rather, either an increase or a decrease of its *degree of confirmation*” (Carnap, 1994: 292). Carnap is explicit that “pragmatics is the basis of all linguistics” (Carnap, 1942: 13). According to Carnap, the central subject matter in pragmatics is the speaker who speaks, hears, or writes the expressions of the language, and the method that one may employ in this field is “entirely empirical” (Carnap, 1939: 4–9). The descriptive nature of the pragmatic concepts is what distinguishes them from other concepts, which are of a more theoretical nature.

Naming, for example, at this stage, is primarily of an indexical or ostensive nature (or simply observational), and in consideration of sense data. Thus the truths regarding linguistic phrases of these sorts are to be considered as special kinds of truth called “factual truths” (F-truth). This means it has to be established via observation, empirical factors, and immediate confirmation of the language community. As mentioned by Carnap, pragmatics is where we test our scientific theories about the actual world or where we start to make new ones (Carnap, 1994).

In general, Carnap considers pragmatics as the realm in which we form explicanda. Later on, in pure semantics, we are to provide explicata for them (Carnap, 1955a: 34). Therefore, the construction of the meaning or *intension* of the terms should start at the pragmatic level. The following is an example.

The explicandum “belief” is considered to be the relationship T (not B), between a person and a sentence (not a proposition); because the relationship B , between a person and a proposition is *nonpragmatical* in the sense that “characterizes a state of a person not necessarily involving language” (Carnap, 1955b: 90). That is to say not a relation of the form

$$B(X, t, p)$$

that would say that the person X at the time t believes that p . But, a relation of the form

$$T(X, t, S, L)$$

that would say that the person X at the time t takes the sentence S of the language L to be true (consciously or not). “Now the *pragmatical* concept of *intension* serves as a connecting link between B and T . Let a relation of the form

$$Int(p, S, L, X, t)$$

say that the proposition p is the intension of the sentence S in the language L for X at t ” (Carnap, 1955b: 90–91)

Once a natural language becomes actualized or activated at the pragmatic level, we may disregard the speaker-world relationship, and go up to the semantics where the designation relationship is our central focus. “If we abstract from the user of the language and analyze only the expressions and their designata, we are in the field of semantics” (Carnap, 1942: 9)

3.2 Semantics

In semantics we disregard the speaker of the language and we will only consider the relation of designation that is the relation between a term and its “designatum” (an inside-language relation). Here is where we assign names, properties, relations, etc. to objects, and indirectly determine the truth conditions of the sentences. The more precise the rules we set up for designation, the more accurate the results (or way of speaking). This accuracy, in turn, leads to less controversy in discourses within the language community. Although we ourselves set up the rules for deciding what is right or wrong according to the system (since we are the ones who are making the conventions), the rules are not arbitrary. They are bound to the empirical node mentioned above. This is explicitly clear from the following quotation where Carnap

is talking about an imaginary language "B" which belongs to the world of facts, and our own established semantics for this language, "B-S", and which has all and only the properties that we have constructed by our rules.

Nevertheless, we construct B-S *not arbitrarily* but with regard to the facts about B. Then we may make the empirical statement that the language B is to a certain degree in accordance with the system B-S. The previously mentioned *pragmatical facts are the basis* [...] of some of the rules to be given later. (Carnap, 1939: 7) (Emphasis mine)

Since the main goal of setting semantic rules is to achieve the highest degree of accordance with facts, we are bound to this accordance, and preferring one semantic system over another is not a mere matter of terminological choice but rather a matter of degree of confirmation with respect to the facts. Here is, in semantics, where we define synonymy and where we form our theories of *meaning*.

Semantics would ideally give us an "interpretation" of the language by which we would be able to understand expressions of the language. According to Carnap (Carnap, 1939: 11), understanding a language, a sign, an expression, or a sentence are all due to the semantic rules of the language system.

Let us not forget that we are not entirely unconcerned with empirical observations (at least as far as it concerns descriptive semantics). But at a certain point when setting up semantic rules of designation, we are no longer concerned with non-linguistic objects. Once a natural language becomes actualized or activated at the pragmatic level, we may disregard the world-speaker relationship, and go up to the semantics where the designation relationship is at the center of attention. Here, naming, for example, has a referential characteristic as opposed to an observational or ostensive characteristic it has in pragmatics. That is to say, in semantic, the word "red", for instance, is considered to be a term (an elementary term) and not sense data; whereas in pragmatics the use of the same word is in consideration of the sense data that would allow its attribution be followed by immediate confirmation (or infirmation) of the language community.

Semantics, according to Carnap, is the lowest level of abstraction. Abstraction in semantics may begin by simply switching our observational concern to our concerns about the occurrences of signs. This switch of attention means nothing more than disregarding empirical factors involved in observation and just focusing on the designation relation between the signs and their designata regardless of their actual existence. At this point we are ready to study the inherited language, built up at pragmatics, as an object by itself; we may call it the "object language". So, the mark for entering into the realm of abstraction is just switching our attention from observation to designation by presupposing the existence of the involved objects (events, relations, etc.); this is very similar to the definition of constructivism i.e., the strict interpretation of "there exists" as "we can construct" (Bridges & Palmgren, 2013). Just as we disregarded empirical factors in observation to focus on the designation relation, we may continue disregarding the factual content of the statements even further in order to ascend to higher abstract levels. Now, we are at the level that is called "pure semantics" (L-semantics; L stands for "logical") (Carnap, 1939). In this special semantics, the designata of the signs (sentences, names, connectives and the like) are not outside of the language system, and they are with regard to solely inside-language elements (e.g., L-implication, L-equivalence, L-true, L-false and the like). Thus, in L-semantics, the truth about atomic and molecular sentences

(L-truth) can solely be investigated via the rules of our conventional truth-value assignments regarding the logical connectives.

Investigating the rules that would allow us to make such truth-value assignments in L-semantics (i. e., assigning L-true or L-false) is the goal of the final part of our language analysis, i. e., the syntax. Now we have passed the skin (pragmatics) and the muscles (semantics), and have reached the skeleton of the object-language (syntax).

3.3 Syntax

In syntax, the relation of designation will be completely disregarded. Here, by formalizing, in a meta-language, we determine and set up the rules according to which we may assign semantic terms such as L-true, L-false, and the like, to sentences. Syntactical rules would serve two purposes: constructing proofs and making derivations⁶. Carnap defines C-true sentences (C for calculus) as “the sentences to which the proofs lead” (Carnap, 1939: 17). Logic is a discipline that takes care of this purpose, and Carnap sees it as a system that has been established and developed by thinkers like Aristotle and Euclid, grown up in the hands of philosophers like Leibniz and Boole, and became more comprehensive by mathematicians and philosophers like Schroeder, Frege, Peano, Whitehead, and Russell, and benefitted a good deal from Hilbert’s axiomatic method (Carnap, 1939: 17)⁷.

At the syntactic level our concerns are no longer the objects themselves (i. e., what they do designate, hence their soundness) but the validity of the structure (or sequentiality) of the objects (or signs). “The syntax of a language, or of any other calculus, is concerned, in general, with the structures of possible serial orders (of a definite kind) of any elements whatsoever” (Carnap, 1937: 6). In propositional logic, we call these structures “rules of inference”. With *modus ponens*, for example, successive true appearances of a material conditional and its antecedent guarantee the true appearance of its consequent. For Carnap, semantic, in general, is an interpretation (true or false)⁸ of a calculus (syntax). That is to say the question of C-truth is all about consistency: “A calculus may (but usually does not) also contain rules which determine certain sentences as C-false. If the rules of a calculus determine some sentence as both C-true and C-false, the calculus is called *inconsistent*; otherwise *consistent*” (Carnap, 1939: 17).

None of the rules of calculus (neither rules of formation nor the rules of transformation) in any way refer to designata, according to Carnap (Ibid: 19). Nevertheless, they have been chosen with regard to the semantic so that the extension of the “C-true”, “C-false”, and “C-implicate” in the syntax coincides with that of “L-true”, “L-false”, and “L-implicate”, respectively, semantic (Ibid). Carnap reminds us that, in principle, we are free to choose from infinite possibility of the rules of calculus; whether or not they are practically justified is another issue:

There are an infinite number of other possible choices of primitive sentences and rules of inference which would lead to the same result. This result gives the practical justification for our choice of the rules of B-C [(the calculus of the language B)]. A calculus in itself needs no justification. (Carnap, 1939: 19–20)

⁶“A derivation leads from any not necessarily C-true sentences, called the premises, to a sentence, called the conclusion [(C-implicate)]” (Carnap, 1939: 17). Proofs could be construed as a special sub-class of derivations, namely ones that proceeded from truths, whereas derivations are any move in the proof system, which might proceed from false premises. The conclusion of a proof is a truth. The conclusion of a derivation is indeterminate.

⁷I should notify that I intentionally limited the discussion here to the first order propositional logic to make my point. One of the major objectives of this paper is to give a general schematic view of Carnap’s LF in order to provide a basis for further discussion on the same topic. Consequently, I will avoid getting into more detailed and technical discussions about analyticity or syntactical rules, and leave that for future papers.

⁸see (Carnap 1939: 21) for the conditions of true interpretation.

As in the case of semantics, in the case of syntax, too, Carnap distinguishes descriptive syntax from pure syntax. "Descriptive syntax is related to pure syntax as physical geometry to pure mathematical geometry; it is concerned with the syntactical properties and relations of empirically given expressions (for example, with the sentences of a particular book)" (Carnap, 1937: 7). Therefore, pure syntax inherits at least some of the properties of the descriptive syntax (if we consider a bottom-up move). Or, pure syntax should be respectful (or loyal) to some descriptive properties by making it possible to provide a useful interpretation (if we consider a top-down move). The relation between descriptive and pure syntax can be defined by introducing "correlative definitions" by means of which "the kinds of objects corresponding to the different kinds of syntactical elements are determined (for instance, material bodies consisting of printers' ink of the form ' $\vee$ ' shall serve as disjunction symbols)" (Ibid). For instance, sentences like "the second and forth sentences of a particular series of sentences (or a passage) contradict one another" or "the third sentence is not syntactically correct (let's say according to English grammar)", are sentences of descriptive syntax. But, sentences like "the sequence $\varphi \supset \psi$ has a general form of $Var(x) Con(x') Var(x'')$ ", where *Var* stands for variable and *Con* for constant, belong to pure syntax. At the same time $Var(a) Con(a') Var(a'')$ still have a descriptive nature. "Pure syntax is thus wholly analytic, and is nothing more than *combinatorial analysis*, or, in other words, the geometry of finite, discrete, serial structures of a particular kind" (Ibid).

When we say that pure syntax is concerned with the forms of sentences, this 'concerned with' is intended in the figurative sense. An analytic sentence is not actually 'concerned with' anything, in the way that an empirical sentence is; for the analytic sentence is without content. The figurative 'concerned with' is intended here in the same sense in which arithmetic is said to be concerned with numbers, or pure geometry to be concerned with geometrical constructions. (Carnap, 1937: 7)

As we saw, pure syntax is the level that completely disregards factual content, and so is maximally conventional⁹. According to this schematic, abstraction could be construed as a bottom-up process of simultaneously disregarding factual content and becoming increasingly conventional. From this point of view, one could see, in general, how abstraction could be subjected to degradation and how it could be correlated with some sort of gradual disengagement process at each step. In order to go from a lower level of abstraction to a higher one, we would disregard a relationship, an object or a predicate of some sort, and make some presuppositions at each step. We also saw in this disengagement process that there is a voluntary element of choice or switch of attention involved (that can be justified pragmatically). This choice may be considered either positively, as to which relationship we want to preserve, or, negatively, as to which relationship we no longer want to be engaged with. One noteworthy observation to make in the picture that Carnap draws of abstraction is to note where the major steps of abstraction are taking place, i. e., from pragmatics to semantics and from semantics to syntax. In both cases, there is a single relationship that is being disregarded. Simultaneously, there are presuppositions to be made regarding the relationship on which we want to concentrate. For example, in the case of moving from pragmatics to semantics, the relationship we wanted to concentrate on was the designation relationship between the signs and their designata, and the relationship that we wanted to disregard was the speaker-world relationship or the relationship between the sign and the actual object; therefore, we presupposed the existence of all designata. In the

⁹Strictly speaking, from lacking content to being conventional, is a non sequitur. While becoming conventional via losing the content might not be the case for some Platonic entities, it is clearly the case for non-Platonist logical empiricists.

next major shift in abstraction from semantics to syntax, we wanted to find valid structures regardless of the designation of their elements; therefore, we presupposed the semantical truth of those elements (i. e., we presuppose the designation relationship holds for all the elements).

In the abstraction model just described, we started the construction of our language system from pragmatics all the way to syntax. According to Carnap, as we will see in the next section, this is only one of the two possible ways of constructing a language system, which we may call a bottom-up method (or an abstractive method). The inverse top-down method (or interpretive method) is also possible, which will be explained in the following section.

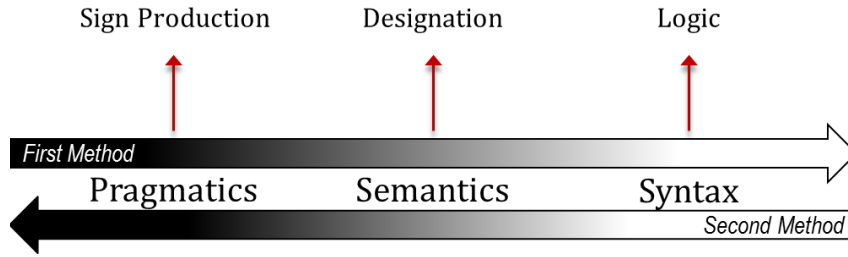


Figure 3.1: Components of a complete language analysis

4 LF and the two methods

Carnap acknowledges that the difference between these three parts is their level of abstraction.

We distinguished three factors in the functioning of language: the activities of the speaking and listening persons, the designata, and the expressions of the language. We abstracted from the first factor and thereby came from pragmatics to semantics. Now we shall abstract from the second factor also and thus proceed from semantics to syntax. (Carnap, 1939: 16)

One may realize that what is interesting here is that Carnap, by establishing the ladder of gradual abstraction (i. e., the gradual loss of factual content), is indirectly suggesting the possibility of a systematic way for dealing with the concept of abstraction. Carnap is clear that if we are to construct a language for science we ought to give up absolute verifiability and consider “gradual confirmation” (Carnap, 1938). He recognizes two methods for constructing a language for science (or basically any sort of language):

Let us suppose we are going to construct an empirical language for the whole of science, [...] At which point in the system of terms shall we begin with the construction? At the one end of the system there are the elementary, concrete terms like ‘blue’ and ‘hard’, which can be applied on the basis of simple observations. On the other end there are the abstract terms as they occur in the most general laws of theoretical physics, e.g. ‘electric field’. There are now two possible ways open to us, each of them having certain advantages. (Carnap, 1938: 33–34)

Before we get into the descriptions of these methods let’s once again consider LF in the following presentation, but this time with respect to the levels of abstraction:

One important point is that, in terms of the factual contents of the sentences, there is some sort of heterogeneity (or factual-conventional duality, if you wish) involved in constructing languages according to this model. That is, the statements in the middle of the factual-conventional spectrum are neither completely factual nor completely conventional. As we have noticed, sentences formed at the lowest level have maximum factual content, and as we go up the abstraction

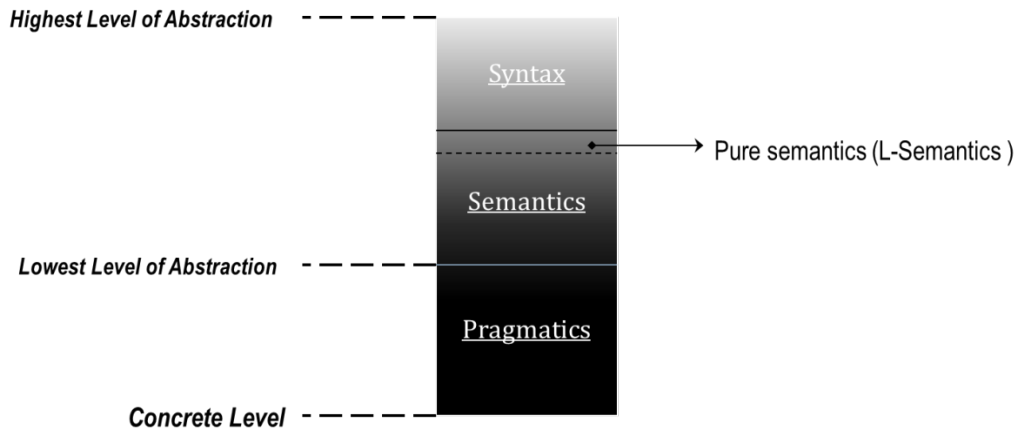


Figure 4.1: Levels of abstraction in a LF

ladder, they lose factual content and become more and more conventional. Consider, for example, how the following set of sentences become more conventional as we go up the abstraction ladder. Looking at the following example gives us a sense of how the statements gradually lose their factual content.

- This is an apple. (Factual)
- The apple is red.
- Red is a color.
- Color is a concept.
- Concept is $F(x)$.
- $F(x)$ is P .
- P is F -determinate.
- P is F -determinate if and only if " $P \wedge \sim P$ " is L -determinate.
- " $P \wedge \sim P$ " is L -determinate if and only if " $P \vee \sim P$ " is C -true. (Conventional)

As we may realize, the construction of a calculus upon which we consider $P \wedge \sim P$ as false (or more specifically, L -false) is purely conventional without any participating factual component. "Now consider the predicator $H \ \& \ \sim H$. No factual knowledge is needed for recognizing that this predicator cannot possibly be exemplified" (Carnap, 1956: 21). In the same way, taking $P \vee \sim P$ as L -determinate (hence analytic) or L -indeterminate (hence synthetic) is entirely based upon the decision of the framework constructor, regardless of any fact. Carnap acknowledges the heterogeneity of LF with respect to the factual content in ESO as well as in other places (e.g., Carnap 1936; 1965). Now we can easily see how we may continue losing factual content up to the syntactical level, where the realm of pure conventions begins.

4.1 The first method

In the first method, we start constructing our language system (LF) by taking elementary terms¹⁰ (such as "blue", "hot", "hard") as primitive terms and then introducing them to higher levels of abstraction. "If a suitable set of elementary terms is chosen as a basis, every other term

¹⁰It is important to notice that regardless of how we arrive at the concepts such as "red", "cold", "hard" and so on (and setting aside the world-language relation), we may still treat them as linguistic entities (belonging to the world of language; where we consider language as an object itself) and call them "elementary terms" (Carnap, 1939: 61). In this way, we pre-assume a certain world-language relation (undetermined and under investigation) about which we are not going to talk, rather we want to talk about language in isolation and as an object in the world. Carnap is clear that "bright", "dark", "red", and other concepts of this sort are "elementary terms" and "meant as properties of

of the language [...] is either definable or at least reducible to them”(Carnap, 1938: 34). The advantage of the first method, according to Carnap, is that “it allows a closer check-up with respect to the empirical character of the language of science. By beginning our construction at the bottom, we see more easily whether and how each term proposed for introduction is connected with possible observations” (Ibid). Thus, the first method is essentially a bottom-up method.

One of the points to which we should pay special attention to, again, is that in the first method of constructing a LF, we are not completely arbitrary precisely because we are empirically constrained. Not paying attention to this point has led to some confusion in the literature. For example, some philosophers, e. g., (Maddy, 2007: 86), hold the idea that making scientific theories is just a mere terminological choice or just a matter of language, for Carnap. As we saw in section 2.2, semantical rules cannot be chosen arbitrarily, and Carnap is clear that they are empirically constrained by factual observations in pragmatics. Since the same relationship that holds between pragmatics and semantics also holds between semantics and syntax (semantics is an abstraction of pragmatics and syntax is an abstraction of semantics), we may say that by the first method of construction, the entire LF is committed to factual observations, and therefore constructing a LF by the first method is not completely arbitrary. Carnap is fairly clear that, in the first method, pragmatic and empirical criteria can be regarded as “practical guides” (or constraints) in setting up rules or making conventions (Carnap, 1939: 6). So in constructing a language system, our choices of rules, for an already-interpreted language (a natural language), are not completely arbitrary. Nevertheless, “nobody doubts that the rules of a pure calculus, *without* regard to any interpretation, can be chosen arbitrarily” (Carnap, 1939: 27) (Emphasis mine).

In sections 11 and 12 of (Carnap, 1939), Carnap is quite clear that in the case of constructing a syntax (or a calculus) for an existing language, which is an instance of employing the first method, we are not completely free and we do bring some commitments to bear. Indeed, we are limited in “some essential respects”, because the syntax must be constructed in such way that it gives us a true interpretation of the existing semantics. The only freedom one may have in this regard would be limited to minor choices in classifying the signs and formulating the rules:¹¹

If a semantical system S is given and a calculus C is to be constructed in accordance with S, we are bound in some respects and free in others. The rules of formation of C are given by S. And in the construction of the rules of transformation we are restricted by the condition that C must be such that S is a true interpretation of C [...]. But this still leaves some range of choice. We may, for instance, decide that the class of C-true sentences is to be only

things, not as sense-data” (Carnap, 1939: 62) thus they already passes a (abstractive) stage that converts sense data into linguistically expressible property-words (i. e., elementary terms); but, while we may consider elementary terms to have independent values, Carnap still consider them as being abstracted from pragmatics. Carnap considers semantic information, in general, to be an approximation to pragmatic information that is achievable by abstraction.

We shall talk about the information carried by a sentence, both by itself and relative to some other sentence or set of sentences, but not about the information which the sender intended to convey by transmitting a certain message nor about the information a receiver obtained from this message. An explication of these usages is of paramount importance, but it is our conviction that the best approach to this explication is through an analysis of the concept of semantic information which, in addition to its being an approximation by abstraction to the full-blooded concept of pragmatic information, may well have its own independent values. (Carnap & Bar-Hillel, 1952: 2–3)

¹¹If we need an example of the choices between different formulations (amongst others), e. g., for propositional logic, we may think of the choices between Łukasiewicz’s system of notations or the notational system of Whitehead and Russell. Both cases, no matter how different they may be, are still committed to satisfying the main condition, which is to provide a true interpretation for the existing semantics.

a proper subclass of the class of L-true sentences, or that it is to coincide with that class or that it is to go beyond that class and comprehend some factual sentences, e. g., some physical laws. [...] This choice, however, is not of essential importance, as it concerns more the form of presentation than the result. If we are concerned with a historically given language, the pragmatical description comes first, and then we may go by abstraction to semantics and to syntax. (Carnap, 1939: 24)

Therefore, in the first method of construction we are not only limited to a true interpretation of the existing semantics, but also committed to the facts of the matter. Carnap also reminds us that the order of the methods is of essential importance because "if we have chosen some rules arbitrarily, we are no longer free in the choice of others" (Carnap, 1939: 25). Then, the first method has an essential priority compared to the second one.

4.2 The second method

Traditionally, being used to the application and rules of one sort of logic might make us prejudiced in favour of that logic; we may even go so far as to construe the system we are familiar with as "obvious". Carnap, on the other hand, sees the possible range of assertions as far more diverse and versatile:

It is important to be aware of the conventional components in the construction of a language system. This view leads to an unprejudiced investigation of the various forms of new logical systems which differ more or less from the customary form (e. g., the intuitionist logic constructed by Brouwer and Heyting, the systems of logic of modalities as constructed by Lewis and others, the systems of plurivalued logic as constructed by Lukasiewicz and Tarski, etc.), and it encourages the construction of further new forms. (Carnap, 1939: 28)

The second method is when we take abstract terms of the highest levels of abstraction or syntax, and introduce them (interpret them) to lower levels all the way to the elementary terms. "If a suitable set is chosen, here again every other term, down to the elementary ones, can be introduced. And here, it seems, explicit definitions will do." (Carnap, 1938: 34). Thus, the second method is a top-down method. The advantage of this method is that "it represents the systematic procedure as it is applied in the most advanced fields of science, especially in physics" (Ibid). If it is to be somewhere, here is precisely where creativity and language planning come to play an essential role.

When using the second method, we are basically free to use whatever calculus (set of syntactical rules) we wish to satisfy our purpose. One of our options is, of course, to stay with the same resulting calculus (let's say classical logic) of the first method and make our changes at lower levels to what Carnap calls "indeterminate statements" (Schilpp, 1963: 920). This might be the most common philosophical/scientific practice, and the result would be LFs sharing the same logic.¹² This fact, of course, does not rule out the other possibility of the adoption of totally different calculi (e. g., intuitionistic logic). If the readjustment¹³ has to be done at highest levels, it will result in a different language. One should keep in mind that

¹²We may think of pure non-Euclidean geometries, which share the same logic as the Euclidean geometry, as an example of this.

¹³In the case of conflict with experience, Carnap distinguishes between two kinds of readjustments (in LF), namely between changing truth-value assignments to the "indeterminate statements" (i. e., statements whose truth value are not fixed by the rules of language, say by the postulates of logic, mathematics, and physics) and changing the language (Schilpp, 1963: 920–921).

even in the case of adopting different calculi, our final interpreted language should ultimately be accountable to the empirical facts of the matter, but the choice of the adoption is only pragmatically, not principally, constrained. There is no logic in choosing logics; one should notice that, in the case of adopting different calculi, we are no longer in the same LF. In the case of changing the language from L_n to L_{n+1} , the concept of “being syntactic”, for example, is totally different in each language. That is, “... is syntactic” in L_n is a different concept than “... is syntactic” in L_{n+1} ; the same is true for “being analytic” (Schilpp, 1963: 920). Therefore, since the property of “being syntactic” (or “being analytic”) is totally dependent on our choice of syntax (which follows no logic and is only justifiable pragmatically), then, the concept of “... is syntactic” is only decidable upon our purely arbitrary chosen calculus. “With respect to a calculus to be constructed there is only a question of expedience or fitness to purposes chosen, but not of correctness” (Carnap, 1939: 25).

The second method of constructing a language system, then, is first to construct a calculus C and then a corresponding semantics S accordingly. And here is how Carnap describes this process:

We begin again with a classification of signs and a system F of syntactical rules of formation, defining ‘sentence in C ’ in a formal way. Then we set up the system C of syntactical rules of transformation, in other words, a formal definition of ‘ C -true’ and ‘ C -implicate’. Since so far nothing has been determined concerning the single signs, we may choose these definitions, i. e., the rules of formation and of transformation, in any way we wish. [...] Then we add to the un-interpreted calculus C an interpretation S . Its function is to determine truth conditions for the sentences of C and thereby to change them from formulas to propositions. [...] Finally we establish the rules for the descriptive sign (Carnap, 1939: 25–26).

The relevance and effectiveness of our choice of C will finally be determined by the richness of the language it yields. Here is where, once again, empirical data will determine how rich and effective the language is for the purpose of communicating among the targeted community.

Now, the question of the conventionality of logic may become clearer. The question, as Carnap puts it (Carnap, 1939: 27), is as follows: are the rules on which logical deduction is based to be chosen at will, and consequently judged only with respect to convenience but not to correctness? Or, is there a distinction between objectively right and objectively wrong systems, so that in constructing a system of rules we are free only in relatively minor respects (as, e. g., the way of formulation) but bound in all essential respects? One may see, by now, that Carnap’s answers to both questions are affirmative. On one hand, in the unobjectionable possibility of constructing a language system from a calculus C to its corresponding semantics S (the second method), we are free in choosing the rules of C and the choice is simply a matter of convenience. On the other hand, in constructing a language system from the point at which the “meaning” of logical signs are given before the rules of deduction are formulated (the first method), the statements might be considered objectively right or wrong on the basis of the presupposed “meaning” of the signs. Carnap summarizes his response to the question of conventionality of logic in the following passage:

Logic or the rules of deduction (in our terminology, the syntactical rules of transformation) can be chosen arbitrarily and hence are conventional if they are taken as the basis of the construction of the language system and if the interpretation of the system is later superimposed. On the other hand, a system

of logic is not a matter of choice, but either right or wrong, if an interpretation of the logical signs is given in advance. But even here, conventions are of fundamental importance; for the basis on which logic is constructed, namely, the interpretation of the logical signs (e.g., by a determination of truth conditions) can be freely chosen¹⁴. (Carnap, 1939: 28)

It is worth emphasizing again that, up to this point, it is fairly evident that the process of losing factual content is a gradual process that coincides with a corresponding gain in conventionality, and that this eventually leads to the pure conventionality of syntax. This point is of special importance later on where we talk about analytic-synthetic distinctions.

5 Confirmation and changes in LFs

The main question in this section is how do LFs differ from one another? When we are to talk about the difference between LFs, one should pay special attention to the essential differences they may have. According to what has been explained so far, the difference between LFs could be construed at two different levels: the difference could be at the syntactic (or abstractive) level or it could be at the semantic (or interpretive) level. When we are considering a syntactic difference, then we are talking about adopting different logical systems (different syntaxes). Hence, one expects a dramatic change in the framework. In that case, we can no longer talk about the concepts of “right” or “wrong”, since they are internal concepts to each framework. On the other hand, keeping the syntax intact, we may talk about semantic differences between two LFs, and then we may talk about right or wrong *interpretations* (provided our explicandum is unique¹⁵).

If we decide to keep the syntax intact, then what is at stake might be the F-truth of the statements that are to be established by confirmation. We should keep in mind that Carnap does not see any fundamental difference between particular and universal sentences regarding confirmation:

Thus, instead of verification, we may speak here of gradually increasing confirmation of the law. Now a little reflection will lead us to the result that there is no fundamental difference between a universal sentence and a particular sentence with regard to verifiability but only a difference in degree. (Carnap, 1936: 425)

In agreement with Reichenbach, Carnap sees every sentence as a probabilistic sentence subjected to gradual confirmation (Carnap, 1936: 425–427); the higher the level of abstraction, the higher the degree of confirmation. For example, confirming the sentence “the apple in my lunch box is red” requires a lower frequency of supporting instances than “all apples are red”.

The facts do not determine whether the use of a certain expression is right or wrong but only how often it occurs and how often it leads to the effect intended, and the like. A question of right or wrong must always refer to a system of rules. (Carnap, 1939: 6)

I do not intend to talk about Carnap's position on universals and particulars here; what I would like to shed light on is Carnap's avoidance of the terms “right” or “wrong”, generally, in the context of these kinds of changes in LF. Although, using his own vocabulary, one should be allowed to use “F-true” (in the case of confirmation) and “F-false” (in the case of infirmation),

¹⁴Compare a two-valued logic with a many-valued logic, for example.

¹⁵In the case that explicandum is not unique we may have equally right, yet different, interpretations. According to Carnap, this is the case in dealing with the concept of probability: “There are two explicanda, both called ‘probability’: (1) logical or inductive probability (probability1), (2) statistical probability (probability2)”. (Carnap, 1973: 269)

the essential points here are two-fold: one is that in this kind of change, where the syntax is intact, the changes are to be implemented at the lower levels of abstraction, and what is at stake is the subject of confirmation and/or the confirmation method. The second point is the concept of gradual confirmation in accordance with the level of abstraction that may or may not lead to the change of the second kind in the LF.

We have to pay attention to the fact that, considering Carnap's LF, what we refer to as language is slightly different than the ordinary or traditional sense of the word "language". According to what we have seen so far, as long as LFs share the same syntax they are not to be considered as different languages but rather different ways of speaking. In this sense, we no longer refer to English and Persian as different languages, as long as we establish our arguments in both English and Persian according to the same set of rules (e.g., the rules of elementary logic). For Carnap, the same is true for different theories (expressed in the same language) using quantification over two sorts of variables, or only one to cover both ranges, as long as they follow the same logical rules:

Thus our present acceptance of the two more explicit forms of translation is merely an introduction of two ways of speaking; it does by no mean imply the recognition of two separate kinds of entities-properties, on the one hand; classes, on the other.(Carnap, 1956: 17)

What makes a confirmation possible, in a LF, is the part of the LF that makes it possible to drive our predictions (and then test them against the facts). This part, of course, is the syntactical rules of the LF. As long as we keep the logical syntax of a LF intact, we may talk about which theory (or which way of speaking) is F-right/confirmed or F-wrong/infrimed. For, the general concept of wrong or right would be decidable only according to the same syntactical rules.

Changing the syntactic rules is, in principle, possible. In this case, what would the resulting LF look like? By changing syntactical rules we are making a radical change in the logical fabric of the LF, and this is the very structure that holds everything together in a LF. The first things to lose as a result of this kind of change would be the concept of "right" or "wrong". "Now, the task is not to decide which of the different systems is "the right logic", but to examine their formal properties and the possibilities for their interpretation and application in science" (Carnap, 1939: 28). The only things left to decide are going to be pragmatic considerations such as simplicity, fruitfulness, and the like, assuming the new syntax could generate a new and fully interpreted language (an artificial language). Again, that it is only in the case of syntactical changes where we refer to different LFs as different languages; as mentioned earlier, in other cases we consider different LFs as different ways of speaking the same language.

To sum up, changing our LF in response to resolving a conflict with experience (or otherwise) can be done in two different ways: one in which the new LF is communicable to the old LF which shares the same logical fabric (and where the statements are sortable according to their degrees of confirmation); and the second in which the new LF is incommunicable to the old one since it does not share the same logical fabric.

6 Analyticity and Quine's objections¹⁶

Before getting into the more detailed discussion, I will present a general picture of how Quine and Carnap construe our belief system, and how they envision the changes in this system.

Quine's proposal: our belief system has a web-like structure that encompasses all our theories, including our theories of logic and mathematics that constitute the core of the web. The periphery of the web is more susceptible to change according to actual facts than the core is. Any changes to this system ought to be initiated from outside of the web even if the readjustments require some changes at the core. Subsequently, any change in our mathematical or logical theories should be essentially in response to some change in our empirical data.

Carnap's proposal: all our beliefs about the world that are expressible in the form of communicable assertions are subjected to a structured system, which provides them meaning. This system (which can be studied in isolation) has a hierarchical structure that is more susceptible to change, according to the facts of the matter at the bottom, and is less susceptible at the top. Since the susceptibility of the structure is inversely proportional to the factual content of the statements, at some point in the structure, the statements have no factual content. The conflict between the system and the facts can be resolved in two ways: (1) implementing changes from the bottom to the top, or (2) making changes in the none-susceptible part of the hierarchy to the desired effect.

So far, we have established the following:

1. The first method of construction is essentially dependent on and is bound to empirical observations (§4.1). Therefore, as far as the first method is concerned, LF is entirely committed to the facts and empirical considerations since it starts from pragmatics (§3.1). (reserving our minor conventional liberties in notations, classifications of the signs, and formulating the rules)
2. The possibility of using the second method with total disregard to the empirical data is an unobjectionable possibility. (§4.2)
3. Carnap admits that resolving a conflict with experience may or may not require syntactical changes. (§4, first quote)
4. Changing the LF is possible in two different ways (§5): by making new ways of speaking (keeping the syntax intact) or making new languages (changing the syntax).
5. The first method is practically prior to the second one.
6. Syntax is purely conventional as it stands at one end of a factual-conventional spectrum or assertion without any reference to the outside objects. (§4)

In his terminology, Carnap makes use of the terms “factual”, “L-indeterminate”, and “synthetic” to refer to the lower levels of abstraction in a LF. “A sentence is called L-determinate if it is either L-true or L-false; otherwise it is called L-indeterminate or *factual*.” (Carnap, 1956: 7). Accordingly, the terms “theoretical”, “L-determinate”, “syntactic”, and “analytic” are being used to refer to the higher levels of abstraction. It is fairly obvious that these terms are intended to use as directional guides. The terms “synthetic” or “analytic” should be considered as indications of a place in a hierarchy, and not a property of an object. To say “all LFs have

¹⁶It may seem that I have not been charitable enough to Quine in this paper as I am citing Quine much less than Carnap. There are two reasons for this: first, since I am defending Carnap's position, it is obvious that I tend to clarify his position by citing his own works. The second reason is that the core of almost all of Quine's arguments against Carnap's points and positions seem to be similar and turn on proving the centrality and fundamentality (absoluteness) of analytic-synthetic distinction. Since I tend to argue against this centrality and fundamentality, citing various versions of the same claim would be redundant.

synthetic statements and analytic ones” is like saying “all geographical regions have an east part and a west part”; no one objects to the east-west distinction, and, for the same reason, the analytic-synthetic distinction is not objectionable, if one considers it this way.

Now, I come to the heart of the matter of the second part of my concerns to briefly show that none of Quine’s major objections address the main points of Carnap’s theory. Quine, according to the evidence given below, clearly does not share the idea that the terms “synthetic”, “factual”, “analytic”, and “theoretical” are supposed to be considered as relative terms pointing to some location rather than absolute ones pointing to some objects. Quine’s confusion is understandable because it is easy to see how a person’s view would have been considered dogmatic and nonsensical if the person thinks of the east-west distinction as an absolute and fundamental one when distinguishing western provinces from eastern ones, for instance.

The ASD is by no means an absolute distinction for Carnap for the following reasons: first, the ASD is a distinction that depends solely upon our decision on where we separate semantics from syntax (simply on our choice of logic). Carnap is fairly clear about this, as I noted earlier. Considering “ $P \vee \sim P$ ” as an L-determinate sentence (or not) is principally based upon our decision, and what to do with the interpretations of P . It is not the case that nature dictates and forces us to consider “ $P \vee \sim P$ ” as an L-determinate sentence, no matter how this principle is inspired by nature. Second, if the ASD was fundamental for Carnap, one could not see any inter-changeability between analytic to synthetic and vice versa. However, in the following letter to Quine, Carnap clearly acknowledges the possibility of such a change, from “being analytic” to “being synthetic” and vice versa:

The difference between analytic and synthetic is a difference **internal** to two kinds of statements **inside** a given language structure; it has nothing to do with the transition from one language to another. “Analytic” means rather much the same as true in virtue of meaning. Since in changing the logical structure of language everything can be changed, even the meaning assigned to the ‘.’ sign, naturally the same sentence (i. e., the same sequence of words or symbols) can be **analytic in one system and synthetic in another**, which replaces the first at some time. (Creath, 1991: 431) (Emphasis mine)

In the previous sections (see §3) you may have noticed that, in introducing and characterizing a LF, we did not make any reference to the ASD, for we did not have to. We saw (see §2 pp. 7) that, by accepting LD, a LF becomes immediate and that there are good reasons for adopting LD. Then, as Carnap mentions in the above quotation, the ASD becomes an internal difference directly decidable upon the set of rules we prefer to take as our set of syntactical rules. Quine, on the other hand, apparently does see this the other way around. Quine holds to the idea that the ASD is a fundamental and absolute distinction for Carnap, and without which neither LF, nor the external-internal distinction, nor other terms such as “artificial language” or “meaning postulates”, and the like, would be possible to use:

Carnap has recognized that he is able to preserve a double standard for ontological questions and scientific hypotheses only by assuming an **absolute** distinction between the analytic and the synthetic; and I need not say again that this is a distinction which I reject. (Quine, 1951: 43) (Emphasis mine)

Modern empiricism has been conditioned in large part by two dogmas. One is a belief in some **fundamental** cleavage between truths which are analytic, or grounded in meanings independently of matters of fact, and truth which are synthetic, or grounded in fact. (Quine, 1951: 20) (Emphasis mine)

In the following quotes, it is even more apparent that Quine takes the ASD as a dogmatic belief that stems from an unnecessary (and perhaps wrong) ontological difference between the two. For him, the ASD refers to a differentiation among objects and entities rather than relative terms in classification:

One conspicuous consequence of Carnap's belief in this dichotomy may be seen in his attitude toward philosophical issues, e. g. as to what there is. It is only by assuming the cleavage between analytic and synthetic truths that he is able e. g. to declare the problem of universals to be a matter not of theory but of linguistic decision. (Quine, 1960)

Now to determine what entities a given theory presupposes is one thing, and to determine what entities a theory should be allowed to presuppose, what entities there really are, is another. It is especially in the latter connection that Carnap urges the dichotomy which I said I would talk about. (Quine, 1951)

Quine also sees Carnap's external-internal distinction regarding existential questions as on par with, or rather, as based upon the ASD. Quine holds that both distinctions would disappear by our trivial choice of the types of variables involved in our scientific theories:

No more than the distinction between analytic and synthetic is needed in support of Carnap's doctrine that the statements commonly thought of as ontological, viz. statements such as 'There are physical objects,' 'There are classes,' 'There are numbers,' are analytic or contradictory given the language. No more than the distinction between analytic and synthetic is needed in support of his doctrine that the statements commonly thought of as ontological are proper matters of contention only in the form of linguistic proposals. (Quine, 1951a: 71)

Quine fails to acknowledge what we explained above concerning the gradual loss of factual content as we move toward more general laws. Because he thinks of the ASD as such a profound and absolute distinction, everything in Carnap's model seems to fall into some sort of black-or-white schema. For Carnap, on Quine's account, statements are either analytic or synthetic, universally (and regardless of our choice of syntax). And, as we saw above, that is not the case for Carnap at all:

Whether the statement that there are physical objects and the statement that there are black swans should be put on the same side of [Carnap's] dichotomy, or on opposite sides, comes to depend on the rather trivial consideration of whether we use one style of variables or two for physical objects and classes. (Quine, 1951a: 69)

In §5 we saw that Carnap(in a way) already admits of the possibility of choosing one or two types of variables (one variable to range over properties/classes and one to range over objects, or just one to range over both), and we saw that Carnap refers to these choices as two different ways of speaking of *the same language*¹⁷. It should be clear that Quine is missing Carnap's main

¹⁷In the §4 of (Carnap, 1956: 17) Carnap speaks about the triviality of referring to the "properties" and "classes". There, he says that the possible translations of "Scot is human" as "Scot has the property of human" or "Scot belongs to (is an element of) the class Human" have the same logical content (as long as the logic stays the same).

[...] the terms 'property' and 'class' seem unnecessary, since there are forms which avoid those terms [(Scot is human)]. Thus the important question may be raised as to whether semantics could not do entirely without those terms. However, we shall first accept them, so to speak, uncritically, endeavoring merely to make their customary use more exact and consistent. [...] Thus our present acceptance of the two more explicit forms of translation is merely an introduc-

point. It is true that we can change our quantification variables, but in both cases we still keep the syntax intact. Still, this is really not the crucial point. Quine goes on to construe Carnap's external-internal questions as category-subclass questions:

The external questions are the category questions conceived as propounded before the adoption of a given language; and they are, Carnap holds, properly to be construed as questions of the desirability of a given language form. The internal questions comprise the subclass questions and, in addition, the category questions when these are construed as treated within an adopted language as questions having trivially analytic or contradictory answers. (Quine, 1951a: 69)

According to our explanations so far, we may agree with Alspector-Kelly (2001) when he says that "Quine's interpretation has Carnap claiming that a sentence turns analytic when the sortal's scope widens far enough for it to count as a universal word. But Quine was wrong" (Ibid: 106). Nevertheless, Quine insists, again, that Carnap's external-internal distinction (as well as his other distinctions, such as ontological-empirical or theoretical-factual) is constructed upon the meaningless ASD. "If there is no proper distinction between analytic and synthetic, then no basis at all remains for the contrast which Carnap urges between ontological statements and empirical statements of existence." (Quine, 1951a: 71). Once again, here we clearly see that Quine base LF on ASD, while, for Carnap, just the inverse is the case. As explained, we begin to construct a LF on the basis of purely empirical statements (in pragmatics); there is no ontological statement to begin with, thus, the validity of our ontological *conclusions* (which we may arrive at them later on, in the process of constructing the LF via abstraction) eventually (and primarily) rest upon our empirical statement's degree of confirmation¹⁸.

We discussed that all these distinctions can be directly predicated upon the conception of a LF (not vice versa), and that a LF is immediate after accepting LD (see §2). That is to say if we agree that logical truths are true by linguistic convention (LD), then we agree that logic is linguistically based, thus we have to look for it in a language system (LF). So, if we want to reject the distinction, all we have to do is to reject LD and LF. One simply cannot accept LD and reject LF. Emptiness of analytic truths from factual content at the syntactic level was very clear to Carnap as well as to other members of Vienna Circle. Carnap is even surprised why Quine finds it is necessary to elaborate on this point, given the prior agreements in Vienna:

The main point of his [Quine's] criticism seems rather to be that the doctrine is "empty" and "without experimental meaning". With this remark I would certainly agree, and I am surprised that Quine deems it necessary to support this view by detailed arguments. In line with Wittgenstein's basic conception [LD], we agreed in Vienna that one of the main tasks of philosophy is clarification and explication. (Schilpp, 1963: 216)

tion of two ways of speaking; it does by no mean imply the recognition of two separate kinds of entities-properties, on the one hand; classes, on the other. (Ibid)

¹⁸To put this in another way, on Carnap's account, although in order to understand (finding a meaning for) empirical statements such as "Scot is human", it is quite possible to adopt a method (of analysis) by which we analyse (or translate) this phrase as "there exist something that has the property of being human and the name Scot" (there are, of course, other ways possible). And it is also possible that employing this method makes the customary use of the original phrase more exact and consistent. But, prior to employing this method, one could make no claim with respect to its uniqueness, obviousness, universality, and absoluteness, which precedes its application. Thus all such claims (including ontological ones) become secondary to the method's application and only pragmatically justified.

The centrality and importance of LD, for Carnap, is even more evident where, in a reply to one of Quine's criticisms against his view on logical truth (Quine, 1960), Carnap hopes Quine would not regard LD as a false statement, because it is only then that Carnap is in a difficult situation:

He [Quine] himself says soon afterwards: "I do not suggest that the linguistic doctrine is false". I presume that he wants to say that the doctrine is not false. (If so, I wish he had said so) He nowhere says that the doctrine is meaningless [...].(Schilpp, 1963: 916)

Carnap again returns to LD, where Quine regards elementary logic as "obvious", when he notes that: "Every truth of elementary logic is obvious (whatever this really means), or can be made so by some series of individually obvious steps." (Quine, 1960: 353). First, Carnap is not sure whether Quine is talking about factual obviousness or theoretical obviousness. In fact, we may never know what Quine meant because he does not distinguish the two:

I shall sometimes be compelled to discuss Quine's views hypothetically, that is to say, on the basis of presumptions about the meanings of his formulations, because I have not been able to determine their meanings with sufficient clarity. [...] I presume that he does not understand the word "obvious" here in the sense in which someone might say: "it is obvious that I have five fingers on my right hand", but rather in the sense in which the word is used in: "it is obvious that, if there is no righteous man in Sodom, then all men in Sodom are non-righteous". [...] If Quine has this meaning in mind, we are in agreement. (Schilpp, 1963: 915)

Given that Quine is in agreement with the second sense of the word "obvious", and since Quine adds later on that LD "seems to imply nothing that is not already implied by the fact that elementary logic is obvious or can be resolved into obvious steps." (Quine, 1960: 355) Carnap shows that Quine's argument against his view on logical truth can actually be regarded as a proof of LD (Ibid: 916):

1. Elementary logic is obvious.
2. LD "seems to imply nothing that is not already implied by the fact that elementary logic is obvious".
3. Whatever is implied by LD is implied by (1).

Hence, since LD is implied by LD:

4. LD is implied by (1).

Again, we can clearly see the importance of LD for Carnap. Thus, and in accordance with what I have explained so far, the assumption of LF comes to us naturally, and from there one may impose their theory about the LF's properties, functions, and the like. It seems obvious that we may only talk about all the different distinctions, such as factual-conventional, etc., once we already accept there is such a thing called LF. It might be quite clear by now that none of Quine's presented objections can be construed as objections against Carnap's main points.

In short, I may summarize my points as follows:

1. If the first method of construction (or making changes) is the one and only possible method, then:
 - a. LF, as a whole, is essentially committed to the facts of the matter, and
 - b. There is only one direction (bottom-up) for change. And,
 - c. In that case, the ASD is useless and redundant.

2. If the second method is possible, in addition to the first one, then:
 - a. LF, as a whole, is only committed to the facts essentially in one direction and pragmatically in the other direction. And,
 - b. There are two possible ways for changing LF. And,
 - c. In that case, the ASD is a useful labelling convention.
3. The second method is possible.

Therefore, the ASD is a valid distinction, and it should be regarded as a relative distinction with respect to a LF.

As it may be seen, one may find the Carnapian LF's structure, built by the first method, quite similar to the Quineian "web of belief" (and, in my view, it is). As described, Carnapian LF's structure holds the same commitments to the facts as the Quineian model does. We saw that Carnap acknowledges the possibility of a bottom-up change in syntax, and he refers to such changes as "radical alterations". For Quine, as well, syntactical changes play the same essential and radical role, and that is why he puts them at the center of his web of belief to keep them safe from immediate changes (Quine & Ullian, 1978: 134). Quine takes syntactical rules to be on par with other rules, and, when the time comes, they are not immune to change. The same can be said for Carnap. The only thing that Carnap points out, and that Quine dismisses, is that in the event that such a change has occurred, we are no longer speaking the same language. Consequently, the major difference between the two is that, for Quine, the only legitimate move for readjusting and modifying the structure of our language system is from the boundary to the core of the web (in the Quineian model) or from the bottom of the LF to the top (in the Carnapian model). For Carnap, on the other hand, the move in the other direction is equally legitimate. Quine's justification for taking this position, according to the above discussion, is the obviousness of elementary logic (whatever this might mean). On the other hand, the obviousness of elementary logic, for Carnap, is a theoretical obviousness and belongs to the most conventional part of our language. Therefore, if we admit our principal ability to change whatever we accept conventionally, then change at the syntactical level is both possible in principle and legitimate.

Another interesting conclusion that we may draw from our discussion is that, according to Carnap, coexisting theories in different languages (adopting radically different frameworks) is possible. But, for Quine, there is only one valid theory, i.e. "the theory". It is the theory that encompasses all our explanations about the world. This is the reason that I find Quine's position rather conservative and more akin to traditional ways of thinking.

7 Conclusion

In light of Wittgenstein's and Neurath's views on language, Carnap puts LD at the core of his philosophy. By adopting LD he is allowed to assume a language-based logical structure. The only stipulation that Carnap puts forward at this point is that the process of identifying this structure primarily starts from the bottom of an abstraction hierarchy of a *natural* language. That is, we move from pragmatics to semantics and then to syntax. Our investigation of this structure would come to an end at the highest level of abstraction (syntax) where all the statements' factual contents have been stripped. Once the structure is known, we may refer to the whole construction, inclusive of all the three parts, as a LF. A LF can be construed as a factual-conventional hierarchy for making sense of assertions. Statements at the bottom have factual content, and, as we go upward, they gradually become partly factual and partly conventional/theoretical. Finally, we will arrive at a point in the syntactical level where all

statements are purely conventional and devoid of any factual content. Constructing a LF can be done in two ways: from the bottom to the top (the first method) or vice versa (the second method). There is always a possibility of readjusting or modifying the LF by changing our conventions at different theoretical levels. Introducing modifications into a LF at any level lower than a syntactical level will eventually produce different ways of speaking, while at the syntactical level, they will produce different languages. In principle, we are free to make moderate or radical changes to the LF. We might modify our LF in order to: (1) reach a higher degree of confirmation (according to empirical considerations), or (2) make a simpler and more suitable LF (according to pragmatic considerations). Thus, making a theory, according to this model, is either empirically constrained (when employing the first method, and keeping the same syntax) or pragmatically constrained (when employing the second method, and replacing a different syntax).

According to the given interpretation of LF, we saw that all of Carnap's distinctions, including the ASD, directly hinge upon the conception of LF. We also saw, for Carnap, that the ASD is by no means an absolute distinction. It depends entirely on what arrays of symbols we construe as a constituent part of the syntax (and on where we draw the analytic line). The ASD can only be defined according to a known structure (it is internal to LF). The distinction, regardless of a defined structure, is absolutely meaningless.

Again, according to this interpretation of a LF, one may clearly see that, at the very least, some of Quine's objections cannot be defended and do not affect the main points of Carnap's theory. From this angle, we may be in a better position to understand other important philosophical debates such as the Frege-Hilbert debate¹⁹ on the foundations of mathematics. One may find no fundamental difference between their accounts. The difference, instead, might lie in their corresponding levels of abstraction that they prefer to adopt. Frege might be more committed to a semantical level, whereas Hilbert is posing his ideas at the syntactical level.

Briefly, in this paper, we have established that the conception of LF is a fundamental and an unobjectionable concept in Carnap's philosophy; therefore, his external-internal distinction follows almost immediately. What we may refer to as the ASD is mostly concerned with identifying the levels of abstraction in a LF, and not a fundamental distinction. Therefore, it can be seen as a relative or methodological distinction depending on our conventional decision about what is to be included as a synthetic statement (e.g., the law of excluded middle may or may not be considered as a synthetic statement). We also demonstrated that a LF, as a whole, ultimately receives its support from the results of our empirical observations. According to these results, one is quite capable of considering Quine's own established "web of beliefs" system as only one of the many possible examples of Carnap's frameworks. Obviously Carnap would be in a complete agreement with accepting Quine's "web of belief" as *a framework* (as far as it concerns the first method), but he would disagree that this is the one and only possible way of constructing frameworks. Accordingly, one may find Carnap's model for language analysis more fruitful and constructive compared to that of Quine's that, in my view, is more akin to traditional ways of thinking about philosophical problems.

In general, in Carnap's philosophy, one may easily recognise that usually the terms "analytic", "theoretical", and "syntactic" rest on one side of the story (the abstract and purely conventional side), and the terms "synthetic", "factual", and "semantic" rest on the other side (the less-abstract and less-conventional side). Each term in each group is used in order to de-

¹⁹The debate on what would be constituent of a definition for point, line, and surface, for instance. See the correspondence between Frege and Hilbert in (Gabriel, Hermes, Kambartel, Thiel, & Veraart, 1980).

scribe different aspects (or subject matter) in speaking of a LF. And all of them are directly and primarily related to the conception of a LF.

Abbreviations

ASD: Analytic-Synthetic Distinction

ESO: "Empiricism, Semantics and Ontology" (Carnap, 1950)

LF: Linguistic Framework

LD: Linguistic Doctrine of logical truths: logical truths are true by linguistic convention.

References

- Akiba, K. (1995). 'Quine and the Linguistic Doctrine of Logical Truth', *Philosophical Studies* 78(3), 237–256.
- Alspector-Kelly, M. (2001). 'On Quine on Carnap on Ontology', *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition* 102(1), 93–122.
- Arnold, J., & Shapiro, S. (2007). 'Where in the (World Wide) Web of Belief Is the Law of Non-Contradiction?', *Noûs* 41(2), 276–297.
- Awodey, S. (2007). Carnap's Quest for Analyticity: The 'Studies in Semantics', in: Friedman, M., & Creath, R., eds., 'The Cambridge companion to Carnap', Cambridge University Press, Cambridge, pp. 226–247.
- Berge, W. H. (1995). 'Carnap and Translational Indeterminacy', *Synthese* 105, 115–121.
- Bird, G. (1995). 'Carnap and Quine: Internal And External Questions', *Erkenntnis* 42, 41–64.
- Bridges, D., & Palmgren, E. (2013). 'Constructive Mathematics', in: E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Winter 2013 Edition)
- Burgess, J. P. (2004). 'Quine, Analyticity and Philosophy of Mathematics', *The Philosophical Quarterly* 54(214), 38–55.
- Carnap, R. (1936). 'Testability and Meaning', *Philosophy of Science* 3(4), 419–471.
- Carnap, R. (1937). *The logical syntax of language*, Routledge, London.
- Carnap, R. (1938). 'Empiricism and the Language of Science', *Synthese* 3(12), 33–35.
- Carnap, R. (1939). *Foundations of logic and mathematics*, University of Chicago Press, Chicago.
- Carnap, R. (1942). *Introduction to semantics*, Harvard University Press, Cambridge MA.
- Carnap, R. (1950). 'Empiricism, semantics, and ontology', *Revue Internationale De Philosophie* 4(11), 20–40.
- Carnap, R. (1955a). 'Meaning and Synonymy in Natural Languages', *Philosophical Studies* 6(3), 33–47.
- Carnap, R. (1955b). 'On Some Concepts of Pragmatics', *Philosophical Studies* 6(6), 89–91.
- Carnap, R. (1956). *Meaning and necessity; a study in semantics and modal logic*, University of Chicago Press, Chicago.
- Carnap, R. (1959). *Introduction to Semantics and Formalization of Logic*, Harvard University Press, Cambridge MA.
- Carnap, R. (1967). *The logical structure of the world [and] Pseudoproblems in philosophy* (2 ed.), Routledge, London.
- Carnap, R. (1973). 'Notes on Probability and Induction', *Synthese* 25(3–4), 269–298.
- Carnap, R. (1994). 'Science and Analysis of Language', *Vienna Circle Institute Yearbook* 2, 291–294.

- Carnap, R., & Bar-Hillel, Y. (1952). *An Outline of the Theory of Semantic Information*, MIT Press, Cambridge MA.
- Creath, R. (ed.) (1991). *Dear Carnap, Dear Van: The Quine-Carnap Correspondence and Related Work: Edited and with an Introduction by Richard Creath*. University of California Press, Berkeley.
- Friedman, M. (1999). *Reconsidering Logical Positivism*, Cambridge University Press, Cambridge.
- Friedman, M. (2000). *A Parting of the Ways: Carnap, Cassirer, and Heidegger*, Open Court, La Salle Ill..
- Friedman, M. (2009). *Carnap on Theoretical Terms: Structuralism without Metaphysics*, Paper presented at the Theoretical Frameworks and Empirical Underdetermination Workshop, Düsseldorf.
- Friedman, M., & Creath, R. (eds.) (2007). *The Cambridge companion to Carnap*, Cambridge University Press, Cambridge.
- Gabriel, G., Hermes, H., Kambartel, F., Thiel, C., & Veraart, A. (1980). *Philosophical and Mathematical Correspondance* (Translation: P. Long & R. White), Basil Blackwell, Oxford.
- George, A. (2000). 'On Washing the Fur without Wetting It: Quine, Carnap, and Analyticity', *Mind* 109(433), 1–24.
- Grice, H. (1956). 'In Defense of a Dogma', *The Philosophical Review* 65(2), 141–158.
- Haack, S. (1993). 'The Two Faces of Quine's Naturalism', *Synthese* 94(3), 335–356.
- Hempel, C. G. (1973). 'Rudolf Carnap, Logical Empiricist', *Synthese* 25(3–4), 256–268.
- Hillier, S. (2009). 'Mathematics in Science: Carnap versus Quine', *History of Philosophy Quarterly* 26(4), 409–420.
- Hintikka, J. (1992). 'Carnap's Work in the Foundations of Logic and Mathematics in a Historical Perspective', *Synthese* 93, 167–189.
- Koellner, P. (preprint) Carnap on the Foundations of Logic and Mathematics.
- Lavers, G. (2012). 'On the Quinean-analyticity of mathematical propositions', *Philosophical Studies* 195(2), 299–319.
- Maddy, P. (2007). *Second Philosophy: A naturalistic method*, Oxford University Press, New York.
- McDermott, M. (2001). 'Quine's Holism and Functionalist Holism', *Mind* 110(440), 977–1025.
- OGrady, P. (1999). 'Carnap and Two Dogmas of Empiricism', *Philosophy and Phenomenological Research* 59(4), 1015–1027.
- Oberdan, T. J. (1992). 'The Concept of Truth in Carnap's Logical Syntax of Language', *Synthese* 93(1–2), 239–260.
- Peacock, H. (2011). 'Two Kinds of Ontological Commitment', *Philosophical Quarterly* 61(242), 79–104.
- Price, H. (1997). 'Carnap, Quine and the Fate of Metaphysics', *The Electronic Journal of Analytic Philosophy* 5, Spring 1997, <http://ejap.louisiana.edu/EJAP/1997.spring/price976.html>
- Price, H. (2009). Metaphysics after Carnap : the ghost who walks?, in: Chalmers, D., Manley, D. & Wasserman, R. (eds.), 'Metametaphysics: New Essays on the Foundations of Ontology', Oxford University Press, Oxford, 320–346.
- Psillos, S. (2000). 'Carnap, the Ramsey-Sentence and Realistic Empiricism', *Erkenntnis* 52(2), 253–279.

- Psillos, S. (2000). 'Rudolf Carnap's 'Theoretical Concepts in Science'', *Studies in History and Philosophy of Science* 31(1), 151–172.
- Quine, W. V. O. (1951). 'On Carnap's views on ontology', *Philosophical Studies* 2, 65–72.
- Quine, W. V. O. (1960). 'Carnap and logical truth', *Synthese* 12, 350–374.
- Quine, W. V. O., & Ullian, J. S. (1978). *The web of belief* (2 ed.), Random House, New York.
- Quine, W. V. O. (1951). 'Two Dogmas of Empiricism', *The Philosophical Review* 60(1), 20–43.
- Quine, W. V. O. (1951a). 'On Carnap's Views on Ontology', *Philosophical Studies* 2(5), 65–72.
- Quine, W. V. O. (1960). 'Carnap and Logical Truth', *Synthese* 12(4), 350–374.
- Richardson, A. (2007). Carnapian Pragmatism, in: Friedman, M., & Creath, R. (eds.) 'The Cambridge companion to Carnap', Cambridge University Press, Cambridge, pp. 295–315.
- Richardson, A. W. (2003). Rudolf Carnap in: M. J. Brucoli & R. Layman (eds.), 'Dictionary of Literary Biography: American Philosophers before 1950' (Vol. 270), Cengage Gale, pp. 27–42.
- Runggaldier, E. (1984). *Carnap's Early Conventionalism: An Inquiry into the Historical Background of the Vienna Circle*, Editions Rodopi, Amsterdam.
- Schilpp, P. A. (1963). *The philosophy of Rudolf Carnap* (The Library of Living Philosophers, Vol. 11). Open Court, La Salle, Ill.
- Soames, S. (2009). Ontology, analyticity, and meaning: the Quine-Carnap dispute, in: David John Chalmers, David Manley & Ryan Wasserman (eds.), 'Metametaphysics: New Essays on the Foundations of Ontology', Oxford University Press, Oxford, pp. 424–443.
- Sober, E., & Hylton, P. (2000). 'Quine', *Proceedings of the Aristotelian Society, Supplementary Volume* 74, 237–299.
- Stein, H. (1989). 'Yes, but... Some Skeptical Remarks on Realism and Anti-Realism', *Dialectica* 43(1–2), 47–65.
- Tsou, J. Y. (2003). 'The Justification of Concepts in Carnap's Aufbau', *Philosophy of Science* 70(4), 671–689.
- Wittgenstein, L. (1958). *Philosophical investigations*, Basil Blackwell, Oxford.
- Yablo, S. (1998). 'Does Ontology Rest on a Mistake?', *Aristotelian Society: Supplementary Volume* 72, 229–261.

Justifying Practical Reasons

Georg Spielthener

University of Dar es Salaam
Department of Philosophy and Religious Studies
P.O. Box 110135
Dar es Salaam
Tanzania
georg_spielthener@yahoo.de

The concept of practical reason is central to contemporary thought on ethics. According to a widely held view, we are acting well if we act for good reasons. On this viewpoint, reasons are fundamental to ethics (and practical philosophy in general) because something matters only if we have some reason to care about it. In the current literature on practical reasons there is, however, a tendency towards regarding the concept of practical reason as primitive and indefinable (see Parfit 2011; Scanlon 2000). Authors simply state that reasons are considerations that count in favour of acting in some way and assume, or write as if they assume, that this phrase does not stand in need of further clarification. This paper will show that more can (and should) be said about practical reasons.

Since the nature of reasons for acting is not well understood, and the uses of ‘reason’ are many and diverse (see Audi 2001; Hubin 2001; Schroeder 2007), I need to distinguish between different legitimate senses of ‘reason’ in order to set aside the one I shall be dealing with in this paper. It takes little familiarity with philosophical discussions on the concept of reasons for action to know that there are competing theories of *normative* reasons. (In this essay, I have nothing to say about *explanatory* reasons.) A common way of classifying practical reasons is by distinguishing *subjective* and *objective* reasons. On the subjectivists’ account, the ultimate source of reasons for an agent is in the valuations of that agent. We have most reason to do whatever would best fulfil our present desires (or the desires we would have under some specified conditions). It is fair to say, however, that this account has recently attracted considerable critique. Some authors admit that agents have sometimes subjective reasons for acting (e.g., Scanlon 2000; Searle 2001), but they deny the claim that *all* reasons for action are based on desires. Others, most notably Parfit (2011), think that the subjectivist account is fundamentally flawed. They hold that “we have reasons to act in some way only when, and because, what we are doing or trying to achieve is in some way good, or worth achieving” (Parfit, 2011: 3). In other words, it is facts that give us reasons for action, e.g., the fact that some act would give us pleasure. Let me illustrate this distinction by way of a simple example. Since you believe that the liquid in the bottle is water and you want to drink water, it has been claimed that you have a *subjective* reason to drink it but, as it is actually petrol, you have no *objective* reason for drinking it (Lenman, 2009: 4).

In this essay, I will show that neither objectivism nor subjectivism constitute the proper view of the nature of normative reasons for acting. I will argue for an intermediary position between these two extremes. My focus is on practical reasons that we can prospectively or retrospectively give to justify what we are planning to do or what we have already done. That is to say, my primary goal is to explain when we have so-called justificatory reasons for φ -

ing (where ‘ φ ’ stands in for some verb of action or for verb phrases). Perhaps it is easiest to grasp the basic idea of justifying reasons by an example. Searle (2001) holds that “for a long time people had a good reason not to smoke cigarettes [...] without knowing that they had such a reason” (Searle, 2001: 99). On this view, those people had an objective reason for not smoking—the fact that it causes cancer. I do not deny that there is a use of ‘reason’ that allows saying that they had a reason for not smoking, but I wish to emphasize that those people had no *justification* for not smoking. They did not know the dangers of smoking and they had no reason for believing that it is a health hazard. If a person quitted smoking, he could not (retrospectively) justify his action, and a person who did not take up smoking in the first place could not (prospectively) justify his refusal. Justificatory reasons, however, are meant to provide a justification for the actions they are reasons for. As we shall see below, neither subjective nor objective reasons provide such justifications. It should be noted that having a justifying reason for doing something does not imply that the agent does it. We can have reasons for acting without acting on those reasons. But if a person has a justifying practical reason he has a justification *for* doing it.

Against this background, the aim of this article is now simple to state. I try to answer the question, “When does an agent have a justifying reason for action?” The main thesis of this paper is that a piece of practical reasoning gives an agent a justifying reason for acting if he has a reason for the premises of this reasoning and a warrant for holding that they logically support the conclusion. To argue for this thesis, I shall (in Section 1) discuss the components of such reasons. Section 2 presents a principle of closure for justifying reasons and explains two key clauses of this principle. In the last section, I show how my account can avoid the regress problem in practical reasoning.

1 Components of justifying reasons for action

Suppose you are on a mountaineering holiday in the Alps. You want to climb a certain mountain and you believe that there are only two routes: The eastern route that leads to the peak and the southern route, which leads to a glacial lake. Since you prefer standing on the peak to reaching the lake, you prefer taking the eastern route rather than the southern. Many writers in the field of practical reasoning, notably neo-Humeans, hold that, in the circumstances, you have a *pro tanto* reason for choosing the eastern route. In this section, I will show that this view is misguided by outlining a more sophisticated account of what it is for an agent to have a reason for acting.

Before going on to a detailed consideration of justificatory reasons, two clarifications are called for. First, in the interest of keeping matters as simple as possible, I shall here be concerned with reasoning *under certainty* only. Reasoning is said to be under certainty if the reasoner knows, at least for practical purposes, of each of his options what the outcomes of his taking it would be. Certainty is the simplest case of practical reasoning because no probabilities enter. Second, practical reasoning requires a choice. If we have no choice – when we slip off the ladder or when our body is held immobile – then we do not reason what to do. This may appear so obviously true as to be hardly worth saying, but many authors seem to have overlooked this fact, and they refer therefore to “desires” or other monadic valuations when they discuss practical reasons. However, when we have to choose between different options the relevant valuations are *preferences*. They are dyadic (or comparative) valuations. I take the term ‘*a* is preferred to *b*’ to mean that the agent assigns more value to *a* than to *b*.

Reasons for action are provided by practical reasoning. But when are the premises of a piece of reasoning a justificatory reason for its conclusion? On my view, three conditions must be satisfied. In what follows, I discuss them one by one.¹

1.1 Epistemic reasons

Authors often hold that the agent's *beliefs* create a reason for ϕ -ing. For instance, Beardman claims that "if you have an end E, and believe that doing M is a necessary and available means to bring about E, then you have a pro tanto reason to M" (Beardman, 2007: 257). But surely not any belief will give you a reason. If you have good evidence that M is not a means to E but you obstinately refuse to accept this evidence, then you have no justificatory reason to M. This point may be thought to be rather trivial and obvious, but it is worth emphasizing here because it has generally been overlooked by authors who defend the subjectivist account of practical reasons.

Other writers claim that an agent has a reason to ϕ -ing if it is a (necessary) means for achieving one of his ends. On this view, it is *facts* that generate reasons. These facts can be about an agent's desires (i. e., the fact *that* an agent desires x) but they are "typically facts about valuable states of affairs that the action is likely to bring about, or valuable properties that the action itself will instantiate" (Gert, 2009: 319). For example, Parfit holds that "if I enjoy walnuts, this fact gives me a reason to eat them" (Parfit, 2011: 32) and Scanlon states that the fact that a friend likes Indian food is a reason for choosing an Indian restaurant (Scanlon, 2000: 50).²

Examples that state facts as practical reasons are often persuasive because they implicate that these facts are known and that they are valued or disvalued (see the examples above). In addition, when we give reasons for what we have done, we often express them in factual language. (I sold the car because it was old.) It is, however, not difficult to show that facts *as such* do not provide justifying reasons. Suppose you suffer from a rare disease and you can only be cured if you take substance X. Nobody knows this or has a reason to believe it. Quite the contrary, the prevailing view of experts is that the only thing that can help you is taking substance Y. I think it is clear that you have then no *justificatory* reason for taking substance X (even though it is a fact that it would cure you) and that you have a good justification for taking substance Y, even though it will not help you.

Neither mere beliefs nor facts can justify our actions. What we need is *reasons*. If you have a reason for believing that you can only reach the peak if you take the eastern route then, I submit, you have a justification for taking it (given that the other conditions are met). That is to say, I hold that you have neither a justifying reason for performing an action if this action is in fact a necessary means for one of your ends nor when you only believe that it is such a means; and you do not even have such a reason if the action is actually a means *and* you believe that it is one. What is required is a reason for believing (i. e., an epistemic reason) that this action is a means to something you have a reason to value.

A vital point to notice here is that justificatory reasons require only that the agent has a reason for believing that an action is appropriately related to something that he has a reason to value. It is *not* required that he actually believes this. There is a familiar epistemological distinction between so-called *doxastic justification* and *propositional justification*. Roughly, S

¹Compare to this discussion (Spielthener, 2012) on which this section draws.

²Other authors who hold this view include Hubin (2001), Schroeder (2007), Setiya (2007), Williams (2001) and Raz who claims that it is the fact that this medicine will alleviate your pain that provides you with a reason for taking it (Raz, 1978: 3).

has doxastic justification that p if he believes that p and is justified *in* believing p . If S has propositional justification for believing p then it is *not* implied that S actually believes p . In the theory of practical reasons there is a similar distinction. An agent has a justificatory reason for ϕ -ing if he has a justification *for* believing that ϕ -ing is a necessary means for achieving one of his rational ends. I shall use the phrases ‘having a reason for believing p ’, ‘having a justification *for* believing p ’ and ‘being justified *in* believing p ’ interchangeably. By all of them I mean, roughly, that if S were to believe p solely on the basis of his justifying reason then S would justifiably believe p .³

1.2 Valuational reasons

Some authors, notably decision theorists, hold that if some formal conditions are met, any occurrent or dispositional desire can provide a reason for action.⁴ This, however, is in philosophy a minority opinion. Some philosophical authors admit that desires can give agents some practical reasons, albeit not the most important ones (e.g., Scanlon, 2000; Searle, 2001), and others, in particular proponents of the neo-Humean theory, restrict the set of reason-giving desires to a special class of intrinsic desires – e.g., those that the agent is not alienated from and that are not impulses (see e.g., Shemmer, 2007).⁵

But as I have already mentioned, in practical reasoning we need a choice between at least two alternatives. I therefore hold that it is more appropriate to focus on *preferences* (i.e., *comparative* valuations) rather than *desires* (which are *monadic* valuations). This raises the question as to whether preferences can provide justifying reasons for action. Here is a case of the simplest kind. You have to fly to Vienna today, and the only seats you can get are on Lufthansa and on Air France. Since you prefer Lufthansa, many theorists hold that this preference provides you with a reason for taking that flight. But there is a general argument, which I can only sketch here, that neither occurrent nor dispositional preferences provide justificatory reasons. It is logically possible that acting reverses an agent’s preferences. Suppose that an agent knows that due to his unfortunate psychological constitution, whenever he prefers x to y and chooses appropriate means for achieving x , then his preference changes. He then prefers y to x . If he knows this, he has no reason for choosing the means for what he prefers. Given his knowledge, he has rather a justifying reason for choosing the means for what he does *not* prefer.⁶

On the account presented here, we have a reason for ϕ -ing only if we have a justification for the preference that is a component of this reason. To illustrate, let us return to our example at the beginning of this section. It should be clear by now that believing that the eastern route leads to the peak and wanting to scale that peak does not give you a reason for taking the eastern

³A similar explication has been given by Coffman (2006: 258). He explains the notion of good evidence as “evidence that would render justified a belief in p were S to so believe on its basis”. In a slightly different terminology, Audi states that a person has a situational (propositional) justification for believing x “if and only if the person has grounds [reason] for it such that in virtue of believing it on the basis of them the person would be justified in so believing” (Audi, 2001: 243, note 26).

⁴For example, Maurice Allais, a prominent decision theorist, is quoted by Broome as saying, “It cannot be too strongly emphasized that there are no criteria for the rationality of ends as such other than the condition of consistency. Ends are completely arbitrary.” (Broome, 1995: 104-105). Some philosophers have endorsed this view. For instance, Bertrand Russell has held that reason “has nothing whatever to do with the choice of ends,” it only requires choosing the right means “to an end that you wish to achieve” (Russell, 1954: 8).

⁵As common in philosophical usage, these authors take ‘desire’ in a broad sense, in which it is a generic name for a large group of pro-attitudes (including intending, wanting, liking, caring, feeling committed) that an agent can have towards an action, outcome, or any other content of his attitudes.

⁶Please notice that I am assuming here that the agent is not pathological in the sense that whenever he chooses means for achieving y his preference changes again and he then prefers x . And so on, and so on.

route. And even if you have a reason for your beliefs and you prefer the peak to the glacial lake you have no justification for taking the eastern route unless you have a *reason for preferring* to stand on the peak.

Since there is wide agreement among authors in the field of practical reasoning that reasons for action require a justification for the valuational element they contain, I shall not expand on this point here.⁷ Leaving details for later (Section 2), I only want to emphasize one implication of my view. In analogy to the epistemic reasons discussed in the previous subsection, having a reason for valuing a state of affairs does not imply that the agent values it. That is to say, we can have a reason for valuing x without in fact valuing it. We often say, and correctly so, that someone should do something even though he does not want to do it (for instance, should exercise despite the fact that he hates it). But we also say that someone should not do something even though he wants to do it (should not smoke although he craves for it). This is consistent with my view because the notions of *preferring x to y* and *having a reason for preferring x to y* are logically independent, and justifying reasons depend on the latter concept only.

Since we can have a reason for doing something without having a preference, my account is not subjective and it is therefore not affected by the critique levelled against subjectivism (see e. g., Parfit, 2011 or Scanlon, 2000).

1.3 Warrant for logical support

Even if an agent has reasons for the premises of a piece of practical reasoning he may still not have a reason for its conclusion. Having a reason for the premises of a valid argument does not imply that a reasoner has also a reason for the conclusion because he may be completely unaware of the logical relationship between the premises and the conclusion. The issue of how justification for the premises of a valid argument is transmitted to its conclusion has been extensively discussed in epistemology with regard to reasons for believing.

Some philosophers hold that we have a reason for the conclusion of a piece of reasoning if we have a reason for its premises and the premises *in fact entail* the conclusion, no matter whether we have a reason for believing that this entailment holds.⁸ There is, however, wide agreement that this view is mistaken. A simple example will suffice to make this clear. A student has a reason for believing that a certain shape is a right triangle. This does not give him a reason for believing that the square on this triangle's hypotenuse is equal to the sum of the squares on its two legs, despite the fact that this is deducible from what the student believes. What the student believes about the triangle provides him only with a reason for holding the conclusion if he has a justification for believing that the conclusion is deducible from his premise-belief. In addition, any set of premises entails many conclusions, which will always include conclusions that are so complex that the reasoner does not even understand them. It is quite implausible to hold that a reasoner is justified in believing all of them just because he is justified in believing the premises (Boghossian, 2001; Feldman, 1995).

Other philosophers take a similarly extreme view by holding that a reasoner only needs to *believe* that the premises entail the conclusion. It is, however, obvious that this view is inadequate. An agent can believe that the premises logically support the conclusion even though they do not support it and there are good reasons for holding that the reasoning is

⁷See, for instance, Korsgaard (1997) and Schroeder (2007). Hubin characterizes this view as holding that “no reasons will be transmitted to the means unless there is a reason for the ends” (Hubin, 2001: 462), and he holds that “everyone should admit that a person has reason to undertake means to those ends she has a reason to bring about” (Ibid: 459).

⁸Among the authors who hold this or a similar view are Boghossian and Williamson (2003) and Cross (2001).

invalid. If you believe $p \rightarrow q$ and q , you may think that these statements entail p . But this does not give you a reason for believing p , even if you have a reason for believing the premises.

Since this view is obviously too weak, some authors prefer formulations that seem to be somewhat stronger. Audi holds that the entailment relation must be “within the scope of one’s understanding” (Audi, 2001: 43^{Page number before or after citation?}). Other writers state that a reasoner must *competently deduce* the conclusion (see McBride, 2014), that he must be *aware* of the logical relation between the premises and the conclusion (Meiland, 1980), that he must *realize* that the premises imply the conclusion (Bogdan, 1985: 55) or that the entailment must be *obvious* (Brueckner, 2000; Coffman, 2006). I think it is fair to say that these formulations are too unclear to be helpful for resolving the issue under consideration.

Some theorists require that a reasoner *knows* that the premises entail the conclusion (Stine, 1976; Okasha, 1999). This may be appropriate if we investigate when the premises of a piece of reasoning guarantee knowledge of the conclusion. But the problem with which I am concerned in this paper is when premises provide a *justification* for the conclusion, and for resolving this problem it is not required that the logical support is known. I therefore agree with Klein (1999) that this requirement is too strong.

What is then required for transmitting the justification for the premises to the conclusion? On my view, an agent needs to be justified in holding that the premises logically support the conclusion. This, I submit, is an intermediate position between the untenable extremes that have been discussed thus far. Be it noted that my account does not require that an agent can present a logical theory that shows that the premises entail the conclusion.⁹ Furthermore, an agent is neither expected to self-consciously draw a conclusion (we are often not fully aware of our reasoning), nor is he required to believe that an entailment relation exists (as said, having a reason for believing p does not imply believing it). My account even permits that the premises do in fact *not* entail the conclusion (one can be justified in believing something that is false). All that is required is that a reasoner has some sort of warrant for holding that, given his premises, it would be irrational to deny the conclusion.

2 Closure for justification in practical reasoning

In line with what I have said in the previous section, I shall now propose a principle of closure for justification that succinctly states the *sufficient* conditions for having a justificatory reason for action. In short, this principle holds that if an agent has a justification for the premises of a piece of practical reasoning and he has a warrant for holding that these premises entail the conclusion of this reasoning, then he has a justifying reason for this conclusion.¹⁰ In other words, I hold that justifying reasons are closed under logical implication.¹¹ Many authors think that we can enlarge what we have a reason for believing by accepting what is entailed by things we have reason to accept. Whether reasons are closed under deduction is an important epistemological question because many sceptical arguments depend on closure. Hence, much has been written about closure principles in theoretical reasoning. In the field of practical

⁹Practical reasoning need not be deductive. In reasoning under uncertainty, the premises provide only probabilistic support for the conclusion. But since I have restricted myself in this paper to reasoning under certainty, I shall concentrate on the deductive case, leaving a treatment of probabilistic reasoning for another occasion.

¹⁰It is important to observe that such reasons are so-called *pro tanto* reasons, i. e., reasons that can be outweighed by better reasons without losing their status as reasons.

¹¹Strictly speaking, closure is a property of sets. Set A is said to be closed under a relation R if every element of this set is such that anything it is R -related to is a member of set A . For instance, the set of statements of a language is closed under finite truth-functional combinations because if we combine statements truth-functionally we get again a statement of this language.

reasoning, however, closure for reasons has (so far as I know) never been discussed, despite its importance for understanding the notion of practical reasons. In what follows, I shall therefore propose and briefly explain a principle of closure for justifying reasons in practical reasoning. Let $P1, P2, \dots, Pn$ be the set of premises of a piece of practical reasoning and C its conclusion.

(CJR) An agent has, at t , a justifying reason for C if he has, at t , a reason for each member of the set $P1, P2, \dots, Pn$ and has, at t , a warrant for believing that this set entails C .

This principle is a material conditional that states the sufficient conditions for justifying reasons.¹² Now I have some comments to make on both clauses of its antecedent.

The second clause of (CJR) states that an agent needs a warrant for believing that the premises logically support the conclusion. Since the logic of practical reasoning is contentious, it will be helpful to outline when, on my view, the premises of a practical argument entail its conclusion.¹³ Like reasoning in general, practical reasoning is *valid* if, and only if, the set consisting of the premises and the negation of the conclusion is *inconsistent*. For definiteness, let me state this basic fact in the following principle of valid practical reasoning.

(P) A piece of practical reasoning that consists of the premises $P1, \dots, Pn$ and the conclusion C is valid *iff* the set $P1, \dots, Pn, \neg C$ is inconsistent.¹⁴

The key notion is now “inconsistency”, which may seem a suspect notion because it is not plain when the premises of a piece of practical reasoning and the negation of its conclusion are inconsistent. My next objective is therefore to explain when practical reasoning is inconsistent.

To understand the concept of *practical inconsistency*, we need to be clear that the premises and the conclusion of a piece of practical reasoning are intentional attitudes (not statements or propositions). In our simple example, the premises are your beliefs that the eastern route leads to the peak, while the southern route leads to the glacial lake and your preference for standing on the peak rather than reaching the lake. The conclusion is your preference for taking the eastern route.¹⁵

But how can *different* intentional attitudes (beliefs and valuations) be inconsistent? Beliefs and preferences have *contents*, which I take to be propositions (for instance, the proposition that you are standing on the peak). The point to emphasize now is that a piece of practical reasoning is *not* inconsistent because its contents are inconsistent in the sense that it is impossible for all of them to be true. What renders practical reasoning inconsistent is rather a special *logical relationship* between the contents of its premises and the content of the conclusion.

¹²Some authors propose closure principles that are *strict* conditionals. But these are unnecessarily strong versions and most philosophers formulate closure principles that are material conditionals.

¹³That practical reasoning can be valid is by no means beyond dispute. Many logicians and philosophers endorse it (e.g., Kenny, 1978 or Broome, 2001), but several writers have argued against it (e.g., Mitchell, 1990 or Searle, 2001). Since a consideration of this issue would take us beyond the confines of the present work and I have argued for the validity of practical arguments elsewhere (see Spielthener, 2007) I will not pursue this issue further here.

¹⁴This principle has been expressed in different ways. Richard Hare contends that “he who assents to the premises is compelled not to dissent from the conclusion, on pain of logical inconsistency” (see Kenny, 1978: 75). Searle holds that the acceptance of the premises of a valid practical argument “commits one to the acceptance of the conclusion” (Searle, 2001: 241). According to Gensler, inconsistency of the premises and the conclusion means that we ought not to combine accepting the premises with accepting the conclusion (Gensler, 1996: 16); and von Wright defined *entailment* between norms as follows: “A consistent set of norms entails a further norm if, and only if, adding to the set the negation norm of this further norm makes the set inconsistent” (Wright, 1999: 34).

¹⁵At this point I envisage the objection that on this view it is not any longer clear how practical reasoning can provide reasons for *action*. In essentials, my reply is very simple and runs as follows. If we have a *direct* reason for preferring ϕ -ing to ψ -ing then we have a *derivative* reason for ϕ -ing, provided that ϕ -ing is an alternative (i.e., can be done by the agent if he chooses to do it). Similar views have been held by Broome (1999) and Searle (2001).

Leaving technicalities aside, I shall now try to elucidate this relationship by referring to our example. (i) Believing that you will stand on the peak if you take the eastern route, (ii) believing that you get to the lake if you take the southern route, (iii) preferring to stand on the peak and at the same time (iv) preferring to take the southern route is practically inconsistent. You would prefer doing something that logically entails an outcome you disprefer and you would disprefer performing an action that entails an outcome you prefer. I think upon reflection it is clear that a reasoner who holds (i) to (iv) is in a condition of mental incoherence that is analogous to theoretical inconsistency. Our example is a piece of valid practical reasoning because, given the premises, preferring not to take the eastern route (i. e., taking the southern route) is practically inconsistent.

The first clause of (CJR) states that an agent must have a reason for the premises. I have nothing to say here about the justification for the beliefs, which is a central problem of epistemology. But I wish to briefly explain how to give a justifying reason for preferences. This issue has received remarkably little philosophical discussion, although it can be said without exaggeration that the assessment of reasons for valuations is an essential component of the general theory of practical reasons.

According to (CJR), you have a reason for taking the eastern route if you have reasons (i) for your beliefs that the two routes lead to the peak and the lake respectively, (ii) for preferring the peak, and (iii) for holding that the premises logically support the conclusion. Call this R_1 .

Suppose now that you are asked why you prefer the peak to the lake (which is here a query for justification not for explanation). Your reasoning for this *preference* has the same structure as your reasoning for the *action* of taking the eastern route. There is no mystery here because the conclusion of both pieces of reasoning is a preference. Let us assume you argue as follows: From the peak I can see a lush valley which I prefer to the lake scenery. If you have a reason for your premises and you are justified in holding that these premises entail the conclusion then you have a justifying reason for preferring the peak. Call it R_2 . You have R_1 if you have R_2 .

Assume you are pressed further. You are asked why you prefer the valley to the lake scenery. If you argue that you enjoy watching the valley more than viewing the scenery at the lake and you have the required justifications then you have given a reason for this preference. Call it R_3 . That is, you have R_1 if you have R_2 and you have R_2 if you have R_3 ; and it seems that this chain of justifications has no end. One may object, therefore, that my account of practical reasons leads to an infinite regress. I will address this issue in the next section.

3 The regress of reasons problem

The regress problem in practical reasoning is fairly simply put. It begins with some expressed valuation, e. g., “I want to buy sports shoes” followed by the question, “Why do you want to buy them?” The regress continues: *Because I want to exercise*. “And why do you want to exercise?” *Because I want to reduce my blood pressure*. “And why do you want to reduce it?” *Because I want to prevent a stroke*, and so forth. Since I hold that only reasons can provide reasons for a valuation or action, the structure of reasoning that I have outlined in the previous sections seems to give rise to such a (potentially infinite) regress. On this view, if you have a reason (R_1) for the conclusion C this reason includes a preference for which you also need a reason (R_2). This reason includes another preference for which you need a further reason (R_3), and so forth. That is, there seems to be a regress from C to R_1 , to R_2 , to R_3 , etc.

But let us have a closer look at this issue. Since (CJR) is a conditional, my account does *not entail* an infinite chain of reasons. That is, it does not follow that if you have a reason for C you have R_1 and if you have R_1 you have R_2 (and thus if you have a reason for C you have

R_2), and so forth. More importantly (and also due to the logical structure of the proposed principle), it is not entailed that if you do *not* have R_2 you do *not* have R_1 and thus you have *no* reason for C . But since the structure of justification is inferential, my account seems to imply that we have only *conditional reasons*. When we give a reason for C by appealing to R_1 , we have not yet shown that C is justified. We have only shown that it is justified *if* the components of R_1 are justified. The justification for C is conditional only. We have a reason for C if we have R_1 , which we have if we have R_2 , and so on. But if all justification is conditional in this sense, then we can never claim that an act is non-conditionally justified.

Authors distinguish three main types of justificatory theories that suggest solutions to this regress problem. *Coherence theories* allow a circling back upon valuations that have already been used in the justificatory chain. But many authors think that any such circularity is vicious (e.g., Klein, 1999). That is to say, it is commonly claimed that the chain of inferential justification must come to an end without circularity. If justification does not end somewhere it seems that our reasons are, in the end, ungrounded. This is the view of foundationalists.

Foundationalism is roughly the view that inferentially justified valuations (or beliefs) are based on foundational valuations (or beliefs) that cannot be justified by further reasons but can nevertheless provide reasons for action.¹⁶ That is, the foundationalist strategy for responding to the regress problem consists in denying that justification can be circular and in claiming that there is a stopping point in the regress of reasons, namely so-called basic valuations (or desires).¹⁷ These valuations are not based on any reasons and hence they are not rationally criticisable. We can, of course, examine the causes and consequences of such valuations but not the reasons on which they are based. There are none. Simple foundationalist views hold that *all* basic desires are reason giving. But this is implausible because whims (e.g., the desire for smashing a malfunctioning machine) and alien desires (i.e., desires that do not reflect an agent's true self; see Hubin, 2003) seem to be clear examples of basic desires that are *not* reason giving. Sophisticated foundationalists concede this, and there have been various attempts to distinguish basic valuations that provide reasons from those that do not (see e.g., Brandt, 1979; Hubin, 2003). But none of them has firmly established itself, and it is fair to say that foundationalist views of justification are still marred by the unresolved problem of basic valuations.

Some epistemologists, notably Klein (1999; 2004), have proposed *infinetism* to avoid the problems of coherentism and foundationalism. According to infinitism, the structure of justifying reasons is infinite because it neither allows circularity nor basic desires as regress stoppers. But also infinitism has been subjected to criticism (see Turri, 2009), and some authors think it is obvious that infinite chains cannot provide justification for beliefs or valuations.

Be that as it may, I think that the outlined theories are mainly of academic interest because under real-world circumstances we use more rough-and-ready procedures of giving reasons. In what follows I shall sketch a common-sense model of valuational justification that can provide a satisfactory solution to the regress problem. (Limitations of space prevent me from defending this view here.)

¹⁶This is not only the view of neo-Humeanism, arguably the most influential theory of practical reasons, but also of non-Humean foundationalists (e.g., Audi, 2001). Foundationalism has, however, attracted much criticism: With regard to *practical* reasoning for example from Parfit (2011) and with regard to *theoretical* reasoning from Klein (1999; 2004), whose critique is relevant to practical reasoning too.

¹⁷An illustrative example of foundationalist thinking was already given by Hume (1777): "Ask a man why he uses exercise; he will answer *because he desires to keep his health*. If you then enquire, *why he desires health*, he will readily reply, *because sickness is painful*. If you push your enquiries farther, and desire a reason *why he hates pain*, it is impossible he can ever give any. This is an ultimate end, and is never referred to any other object" (Hume, 1777/1975: 293).

In everyday practical discourse, chains of justification are neither infinite nor is there any need to go up the chain until we reach so-called basic desires. Our chains of justification tend to be rather short. If Jones tells us that he wants to take up exercise we may want to know his reason for it. If he tells us that he wants to reduce his blood pressure we may still inquire why he has this goal. But if he tells us that lowering his blood pressure is necessary for maintaining his health, there will be *no need* for further reasons.¹⁸ In our daily practical reasoning, we justify our actions (and valuations) to another person until we reach common ground, i. e., until this person *accepts* a consideration as a reason; and the action is then justified relative to this basis. We accept reasons for valuations (and actions) if we agree with the factor that has been given as a reason. For example, we accept keeping one's health as a reason for exercising if we concur (possibly with a reason) with this aim and (possibly justifiedly) think that exercise is a means for preserving health.

The point to emphasize now is that this view has two interesting consequences. First, accepting reasons stops the chain of reasons. It does so not because no further reasons *can* be given but because, in the circumstances, there is *no need* for further reasons. We do not have to postulate basic desires for which no further reasons can be given. Chains of reasons do not end because basic desires have been reached but because, in a given situation, we do not need further reasons. To clarify this point, we can adapt Karl Popper's view about scientific reasoning to our problem of practical justification. On this view, our justifications are like "piles driven down from above into the swamp, but not down to any natural or "given" base; and if we stop driving the piles deeper, it is not because we have reached firm ground. We simply stop when we are satisfied that the piles are firm enough to carry the structure, at least for the time being" (Popper, 1968: 111). Should, however, a need for further justification arise (e. g., because new information becomes available), my account allows adding additional reasons at any time and it even permits that the chain of reasons becomes endless.

Second, accepting reasons also alleviates the problem of conditional reasons (see above). Suppose that Jones (elliptically) argues that he has to take up exercise (C) to reduce his blood pressure (R₁). Smith wants to know why he wants to reduce blood pressure and Jones replies that he wants to prevent a stroke (R₂). If Smith accepts R₂ as a reason for wanting to reduce blood pressure (and accepts that Jones has the required epistemic reasons) then he is on my account *inconsistent* if he does not accept that Jones has a reason for taking up exercise. That is, if Smith accepts R₂, he is *irrational* if he does not accept that Jones has a reason for exercising. To be clear, I do *not* hold that Smith's accepting R₂ entails that Jones *has* a reason for exercising, I rather hold that Smith's accepting R₂ logically commits him *to accept* that Jones has this reason. Smith has to accept this because, on my account, R₂ (together with the epistemic reasons of R₁) logically entails C (and accepting the premises of a valid argument while denying its conclusion is irrational). Jones' reason is therefore not only conditional. This point can be generalized. If A argues for a conclusion and B accepts a reason in the chain of reasons, then A can hold that B must (on pain of inconsistency) accept that A has a reason for the conclusion. This result, I think, is crucial because it shows how to respond to the claim that on the account presented here, we have only conditional reasons. We can provide reasons whenever we are challenged. But once an opponent accepts one of the reasons in our justificatory chain, he

¹⁸Justification is in this respect similar to definition. It may seem that defining is an endless process because the defining terms are themselves in need of definitions. But in practice, chains of definitions end (even if we do not just take some terms as basic). If you do not know what 'vixen' means and you are told that a vixen is a female fox, there is very likely *no need* for further definitions.

must, to avoid inconsistency, accept that we have a reason for our valuation. I don't think that a successful justification of an action requires something deeper.

I believe that this brief exposition of my account of valuational justification, incomplete though it is, can alleviate the philosophical worries engendered by the threat of an infinite regress of justificatory reasons. It allows us to give reasons whenever we need some, without being committed to the controversial claim that some valuations are foundational in the sense that no reasons can be given for them. The account presented here is also preferable to *infinetism* because it does not imply that the justificatory chain is infinite. As noted previously, it is controversial whether infinite chains can justify any valuation. According to my view, the process of reason giving comes to an end, possibly only temporarily, when "the piles are firm enough to carry the structure".

I think my account is also preferable to coherence theories of justification. I have already mentioned that these theories imply that the process of justification loops back upon itself, and that many authors consider this as unacceptable. Nonetheless, some authors have taken coherence theories to be central to justification in practical reasoning (e.g., John Rawls and his followers). The general appeal of coherentist conceptions of justification is that they do not require distinguishing between foundational and derived valuations. Justification is rather a matter of mutual support of valuations and beliefs. Unfortunately, however, authors in the field of practical reasoning have not provided the details necessary for their account to advance beyond the metaphorical stage. Some of the challenges that a coherence theory faces are the following. (i) We need to know when a set of preferences is coherent. But to the best of my knowledge, no plausible proposal concerning the precise definition of coherence of (sets of) preferences has been presented. (ii) Incompatible sets of preferences and beliefs can be equally coherent. If coherence is sufficient for justification then all these incompatible sets will be justified, which thoroughly undermines the plausibility of the coherence theory. (iii) Coherence admits of degrees. That is to say, sets of preferences can be more or less coherent. A coherence-driven theory of justification would require choosing the most coherent set. But according to Milgram, no comparative notion of coherence that is precise enough to give us a clear answer which preferences are most coherent has been developed. He therefore holds that "appeals to coherence are empty and the merits of coherence-driven accounts of practical reasoning cannot be assessed" (Milgram, 2001: 13). The coherence theory may represent an initially suggestive solution to the problem of valuational justification. It proposes a way of thinking about justification as arising from fitting everything together into one coherent view. But I think it is fair to say that currently, there exists no convincing account of coherentist justification in practical reasoning.

For these reasons, I hold that the model of valuational justification presented here is an improvement over its rival accounts. If an agent has provided reasons until common ground has been reached, he has done what we might realistically expect of him.

4 Conclusion

The question I have been addressing in this paper is "When does an agent have a justifying reason for action?" We have arrived at the following results: A piece of practical reasoning gives an agent a reason for action if he has a reason for its premises and a warrant for holding that these premises logically support the conclusion. Roughly, that is to say that justifying reasons are closed under logical implication. Contrary to how it may appear at first sight, this view need not give rise to an infinite regress of reasons. In everyday reasoning, the process of giving reasons ends if there is no need for further reasons because the contestants have

reached common ground. In addition, my account does not imply that justifying reasons are conditional reasons only because if a person accepts one link in a chain of reasons she would be irrational if she denied that the agent has a reason for action.

References

- Audi, R. (2001). *The architecture of reason: The structure and substance of rationality*, Oxford University Press, Oxford.
- Beardman, S. (2007). 'The special status of instrumental reason', *Philosophical Studies* 134, 255–287.
- Bogdan, R. J. (1985). 'Cognition and epistemic closure', *American Philosophical Quarterly* 22(1), 55–63.
- Boghossian, P. (2001). 'How are objective epistemic reasons possible?', *Philosophical Studies* 106, 1–40.
- Boghossian, P. and Williamson, T. (2003). 'Blind reasoning', *Proceedings of the Aristotelian Society. Supplementary Volumes* 77, 225–293.
- Brandt, R. B. (1979). *A theory of the good and the right*, Oxford: Clarendon Press.
- Broome, J. (1995). *Weighing goods: Equality, uncertainty and time*. Oxford: Basil Blackwell.
- Broome, J. (1999). 'Normative requirement', *Ratio* 12, 398–419.
- Broome, J. (2001). 'Normative practical reasoning', *Proceedings of the Aristotelian Society Supplementary Volume* 75(1), 175–193.
- Brueckner, A. (2000). 'Klein on closure and skepticism', *Philosophical Studies* 98(2), 139–151.
- Coffman, E. J. (2006). 'Defending Klein on closure and skepticism', *Synthese* 151(2), 257–272.
- Cross, C. B. (2001). 'The paradox of the knower without epistemic closure', *Mind* 110(438), 319–333.
- Feldman, R. (1995). 'In defence of closure', *The Philosophical Quarterly* 45(181), 487–494.
- Gensler, H. J. (1996). *Formal ethics*, Routledge, London.
- Gert, J. (2009). 'Practical rationality, morality, and purely justificatory reasons', *American Philosophical Quarterly* 37(3), 227–243.
- Hubin, D. C. (2001). 'The groundless normativity of instrumental rationality', *The Journal of Philosophy* 98(9), 445–468.
- Hubin, D. C. (2003). 'Desires, whims and values', *The Journal of Ethics* 7(3), 315–335.
- Hume, D. (1777/1975). *Enquiry concerning the principles of morals* (3rd rev. ed.), L. A. Selby-Bigge (ed.), Clarendon Press, Oxford.
- Kenny, A. J. (1978). Practical reasoning and rational appetite, in: J. Raz (ed.), *Practical reasoning*, Oxford University Press, Oxford, pp. 63–80.
- Klein, P. D. (1999). 'Human knowledge and the infinite regress of reasons', *Noûs* 33, 297–325.
- Klein, P. D. (2004). 'What "is" wrong with foundationalism is that it cannot solve the epistemic regress problem', *Philosophy and Phenomenological Research* 68(1), 166–171.
- Korsgaard, C. M. (1997). The normativity of instrumental reason, in: G. Cullity & B. Gaut (eds.), *Ethics and practical reason*, Clarendon Press, Oxford, pp. 215–254.
- Lenman, J. (2009). Reasons for action: Justification vs. explanation. In E. N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*. <http://www.stanford.edu/entries/reasons-just-vs-expl>, accessed 21 April 2010.
- McBride, M. (2014). 'Sensitivity and closure', *Episteme* 11(2), 181–197.

- Meiland, J. W. (1980). 'What ought we to believe?', *American Philosophical Quarterly* 17(1), 15–24.
- Milgram, E. (2001). Practical reasoning, the current state of the play, *in*: E. Milgram (ed.), *Varieties of practical reasoning*, The MIT Press, Cambridge MA, pp. 1–26.
- Mitchell, D. (1990). 'Validity and practical reasoning', *Philosophy* 65, 477–500.
- Okasha, S. (1999). 'Epistemic justification and deductive closure', *Crítica* 31(92), 37–51.
- Parfit, D. (2011). *On what matters, vol. 1*, (edited and introduced by S. Scheffler), Oxford University Press, Oxford.
- Popper, K. R. (1968). *The logic of scientific discovery*, Hutchinson, London.
- Raz, J. (1978). Reasons for actions, decisions and norms, *in*: J. Raz (ed.), 'Practical reasoning', Oxford University Press, Oxford, pp. 128–143.
- Russell, B. (1954). *Human society in ethics and politics*, Routledge, London.
- Scanlon, T. M. (2000). *What we owe to each other*, Belknap Press, Cambridge MA.
- Schroeder, M. (2007). 'Weighting for a plausible Humean theory of reasons', *Noûs* 41(1), 110–132.
- Searle, J. R. (2001). *Rationality in action*, MIT Press, Cambridge MA.
- Setiya, K. (2007). 'Cognitivism about instrumental reason', *Ethics*, 117, 649–673.
- Shemmer, Y. (2007). 'Desires as reasons', *Philosophy and Phenomenological Research* 75(2), 326–348.
- Spielthener, G. (2007) 'A logic of practical reasoning', *Acta Analytica*, 22(2), 139–153.
- Spielthener, G. (2012) 'Justificatory reasons for action', *The European Journal of Analytic Philosophy*, 8(2), 56–72.
- Turri, J. (2009). 'On the regress argument for infinitism', *Synthese*, 166(1), 157–163.
- Williams, B. (2001). Postscript: Some further notes on internal and external reasons, *in*: E. Millgram (ed.), 'Varieties of practical reasoning', MIT Press, Cambridge MA, pp. 91–97.
- Wright, G. H. v. (1999). Value, norm, and action in my philosophical writings, *in*: G. Meggle (ed.), 'Actions, norms, values: Discussions with Georg Henrik von Wright', de Gruyter, Berlin-New York, pp. 11–33.

O caráter definicional *sui generis* dos predicados tarskianos de verdade

Luciano Vicente

Departamento de Filosofia,
Instituto de Ciências Humanas,
Universidade Federal de Juiz de Fora (ICH-UFJF).
Rua Jos Loureno Kelmer, S/N - So Pedro,
Juiz de Fora - MG, CEP 36036-900
luciano.vicente@ufjf.edu.br

Abstract

The *definitional* feature of Tarski's theory of truth will be the subject of this paper. In fact, addition, subtraction, multiplication and divisibility were well-known mathematical concepts before the accurate Peano (Dedekind) formalization. Analogously, the Tarski's metatheory could be an accurate formalization of '*x* is a formula', '*x* is the reference/sense of *y*' and '*x* is a true sentence', all them introduced by definition. However, '*x* is a true sentence', because of the paradoxes, cannot be an accurate formalization of truth predicate of ordinary language. The question is: which concept of truth does the Tarskian '*x* is a true sentence' formalize? The answer is simple and not new, but its meanders are informative and enlightening.

1 Introdução

O caráter *definicional* e, incidentalmente, o *semântico* da **Teoria Tarskiana da Verdade** serão objeto de discussão nesse ensaio. Não se pretende, contudo, negá-los, mas tão-somente melhor elaborá-los e/ou precisá-los.

De fato, nenhum esforço interpretativo profundo será necessário às nossas pretensões, nossa primeira questão será simplesmente: “Quais os objetivos de Tarski em “O Conceito de Verdade nas Linguagens Formalizadas”?”.

Nossa discussão, porém, estará a quem das sutilezas associadas ao tratamento do tema na literatura filosófica contemporânea, embora nossas conclusões sejam congruentes e, eventualmente, concordantes com muitos dos resultados dessa literatura um tanto mais sutil.

2 Estruturas tarskianas

Uma resposta incompleta e, contudo, relativamente precisa para a questão colocada anteriormente é que Tarski pretende construir uma definição (1) materialmente adequada e (2) formalmente correta do predicado '*x é uma sentença verdadeira*'.

A pergunta estrategicamente crucial para nossos objetivos é algo mais direta: “Como Tarski pretende alcançar seus objetivos?”.

Uma resposta relativamente completa e historicamente mais precisa do que é costume é que Tarski parte de uma (a) uma *ciência dedutiva*¹ *C* (b) definida na *linguagem formalizada*² *L* (c)

¹Em alemão, *deduktive Wissenschaft*; em inglês, *deductive science*.

²Em alemão, *formalisierte Sprache*; em inglês, *formalized language*.

cujas atribuições $\mathcal{A}$ de *sentido/referência*³ às constantes de $\mathcal{L}$ seria compatível com $\mathcal{C}$ para, então, construir **(d)** uma *linguagem formalizada* $\mathcal{L}'$, $\mathcal{L}'$ é a *metalinguagem tarskiana*⁴ de $\mathcal{L}$, **(e)** uma *ciência dedutiva* $\mathcal{C}'$, a *metaciência tarskiana* de $\mathcal{C}$, **(f)** uma *fórmula* $V(x)$ de $\mathcal{L}'$, $V(x)$ é o *predicado tarskiano de verdade* de $\mathcal{L}'$ conforme $\mathcal{A}$, **(g)** uma *função* n das expressões de $\mathcal{L}$ em certos termos de $\mathcal{L}'$, $n(\alpha)$ é o *nome descritivo-estrutural de Tarski*, de $\mathcal{L}$, em $\mathcal{L}'$ e, finalmente, **(h)** uma *função* t das fórmulas de $\mathcal{L}$ em certas fórmulas de $\mathcal{L}'$, $t(\alpha)$ é a *tradução tarskiana* da fórmula α de $\mathcal{L}$ em $\mathcal{L}'$.

Antes de entrarmos em alguns detalhes, notemos, primeiramente, que a questão da correção formal da definição é resolvida imediatamente por meio da construção da fórmula $V(x)$, ' $V(x)$ ' é tão-somente uma abreviação. Para uma definição propriamente dita (e correta formalmente), é necessário tão-somente adicionar uma nova constante à metalinguagem tarskiana $\mathcal{L}'$.

Em outras palavras, Tarski, realmente, propõe algo que é potencialmente uma definição (formalmente correta); entretanto, algumas perguntas, talvez, menos-prezadas nesse contexto são: “Em que contexto essa definição é apresentada?” e “Do que é exatamente essa definição?”. Se há algum valor na discussão proposta nesse ensaio, então as respostas “Na metalinguagem $\mathcal{L}'$ ” e “Da concepção clássica de verdade” serão ambas insuficientes.

Voltemos, estrategicamente, ao problema da adequação material. Tarski julga resolvê-lo estabelecendo o seguinte metateorema:

Teorema I. Para qualquer sentença α de $\mathcal{L}$, $\vdash_{\mathcal{C}'} t(\alpha) \leftrightarrow V(n(\alpha))$.

O Teorema I é, desse modo, um metateorema relativo à estrutura $\langle \mathcal{L}, \mathcal{A}, \mathcal{C} \rangle$ e, segundo Tarski, estabelece que o predicado $V(x)$ é, por assim dizer, factual ou materialmente adequado ou, dito de outro modo, a teoria $\mathcal{C}'$ módulo $V(x)$, t e n é uma teoria materialmente aceitável da propriedade ' x é uma verdade relativamente à atribuição $\mathcal{A}$ de significado à linguagem $\mathcal{L}'$ '.

Algumas definições auxiliares serão úteis nesse ponto.

Definição I. $\langle \mathcal{L}, \mathcal{A}, \mathcal{C} \rangle$ é uma *estrutura tarskiana de partida* se $\mathcal{L}$ é uma linguagem formalizada (no sentido de Tarski), $\mathcal{A}$ é uma atribuição de significado às constantes de $\mathcal{L}$ e $\mathcal{C}$ é uma ciência dedutiva em $\mathcal{L}$ compatível com $\mathcal{A}$.

Definição II. $\mathcal{C}' = \langle \mathcal{L}', \mathcal{C}', V(x), n(x), t(x) \rangle$ é uma *estrutura tarskiana de chegada* relativa à estrutura tarskiana de partida $\mathcal{C} = \langle \mathcal{L}, \mathcal{A}, \mathcal{C} \rangle$ se cada dos elementos de $\mathcal{C}'$ é construído, a partir de $\mathcal{C}$, segundo as especificações de Tarski em “O Conceito de Verdade nas Linguagens Formalizadas”. É possível, agora, reformular o Teorema I:

Teorema I'. Seja a estrutura tarskiana de partida $\langle \mathcal{L}, \mathcal{A}, \mathcal{C} \rangle$ e a construção da estrutura tarskiana de chegada $\langle \mathcal{L}', \mathcal{C}', V(x), n(x), t(x) \rangle$, segue-se que, para qualquer sentença α de $\mathcal{L}$, $\vdash_{\mathcal{C}'} t(\alpha) \leftrightarrow V(n(\alpha))$.

O ponto crucial dessa reformulação é que os elementos das estruturas tarskianas de partida e chegada são *prima facie* apresentados como intrinsecamente interrelacionados; é somente na hipótese de uma estrutura de partida que podemos construir uma estrutura tarskiana de chegada e, além disso, o “significado” do predicado tarskiano de verdade é determinado por meio dos *nomes descritivos-estruturais* e da *tradução tarskiana*. De fato e mais profundamente, como tentarei argumentar, o “significado” dos elementos da estrutura de chegada é determinado apenas na medida em que todos os elementos das estruturas tarskianas são considerados conjuntamente.

³Em alemão, *Sinn/Bedeutung*; em inglês, *sense/meaning*.

⁴A designação é minha.

2.1 Linguagens formalizadas

Uma das diferenças entre as *linguagens formalizadas* de Tarski e as *linguagens formais* de nossos manuais é simplesmente uma questão de *ênfase*: as primeiras são pensadas como resultado de processos específicos de formalização, enquanto as últimas são pensadas como possivelmente realizando diversas interpretações diferentes. Na estrutura tarskiana de partida $\langle \mathcal{L}, \mathcal{A}, \mathcal{C} \rangle$, a atribuição $\mathcal{A}$ de sentido/referência relativa à linguagem $\mathcal{L}$ está, no caso do artigo de Tarski, imediatamente pressuposta e a própria discriminação entre linguagem e atribuição de “significado”, que está implícita em nossa proposta de especificação dos elementos de $\langle \mathcal{L}, \mathcal{A}, \mathcal{C} \rangle$, é um tanto artificial no contexto do artigo de Tarski.

De fato, as diferentes interpretações (em sentido próximo ao da atual da Teoria dos Modelos) ou as diferentes “significações” possíveis de $\mathcal{L}$ são, naquilo que é essencial à construção tarskiana, ignoradas por Tarski. Ou melhor, para Tarski, uma linguagem sem um “significado” associado é somente um conjunto de expressões e não uma linguagem, o significado é critério de identidade da linguagem: “*A mesma expressão pode, em uma linguagem, ser um enunciado verdadeiro e, em outra, um enunciado falso ou uma expressão destituída de significado.*” (Tarski, 2006: 20-21)⁵. Como foi dito, o contraste aqui é entre uma *linguagem formalizada*, ou seja, entre o resultado da formalização de um conteúdo semântico $\mathcal{A}$ específico e uma *linguagem formal* como passível de receber diferentes interpretações⁶.

Em todo caso e isso é importante para nós, seja na perspectiva das linguagens formalizadas de Tarski, seja na perspectiva mais próxima da Teoria dos Modelos, a linguagem da estrutura de partida é sempre acompanhada de “significado”. De fato, a teoria de Tarski pode ser dita *semântica* devido ao papel desses “significados” na construção da estrutura tarskiana de chegada e, portanto, do próprio predicado tarskiano de verdade; nesse sentido, a verdade tarskiana é parasitária do significado⁷.

2.2 Ciências dedutivas

Uma *ciência dedutiva* $\mathcal{C}$ (no sentido do artigo de Tarski) é simplesmente uma axiomatização (possivelmente parcial) do domínio de “significação” referente à atribuição $\mathcal{A}$ feita na linguagem formalizada $\mathcal{L}$ e, portanto, $\mathcal{C}$ deve ser compatível com $\mathcal{A}$. Em outras palavras, as ciências dedutivas de Tarski são essencialmente teorias formais no sentido padrão pós-hibertiano, no caso, formalizações parciais de $\mathcal{A}$.

2.3 Atribuições de “significado”

Como deve ter ficado claro, rigorosamente falando, Tarski não atribui “significado” à linguagem formalizada da estrutura de partida; no caso, o “significado” é a própria teoria informal, sempre pressuposta, associada ao processo de formalização.

Outro ponto interessante é que os termos *Sinn* (sentido) e *Beudeutung* (referência) são usados indiscriminadamente por Tarski no sentido, como dito anteriormente, de “significado subjacente à teoria informal associada ao processo de formalização que resultou na linguagem

⁵“*The same expression can, in one language, be a true statement, in another a false one or a meaningless expression.*” (Tarski, 1956: 153)

⁶É essa possibilidade que permite, entre outras coisas, um estudo da linguagem formal independentemente dessa ou daquela interpretação particular.

⁷Desse ponto de vista, os predicados tarskianos da verdade não deveriam ser tomados como explicação do conceito de significado (pelo menos, não sem modificações); pois, na teoria de Tarski, o “significado” das expressões é tomado como um dado. É o que Kirkham (2001: 178-181), argumentando em outro contexto e de maneira alternativa, sustenta; Davidson em ‘Truth and Meaning’ (2006: 155-170), por sua vez, rearranja completamente os parâmetros da questão, na medida em que pretende explicar o significado pelo recurso ao predicado de verdade.

formalizada em questão”. De modo que ou a distinção fregeana entre *Sinn* e *Bedeutung* é ignorada por Tarski ou é tida como irrelevante ao tipo de teoria formalizada em causa.

2.4 Metalinguagem tarskiana

Diferentemente do que acontece com a linguagem formalizada da estrutura de partida, a questão do “significado” da metalinguagem tarskiana é algo tanto mais complicada quanto interessante (em todo caso, trata-se, como no caso anterior, de uma linguagem formal). De fato, o “significado” é, por assim dizer, construído gradativamente e, além disso, estabelecido pela interrelação complexa entre as constantes da própria metalinguagem e das funções meta-metalinguísticas *nomeação*, *n*, e *tradução*, *t*.

Exceção feita ao Teorema I — pensado como critério de adequação que as possíveis ciências dedutivas que tratam do predicado de verdade deveriam satisfazer (a famosa Convenção T) —, Tarski não discute os critérios abstratos que regulem a adequação das relações complexas entre *metalinguagem*, *nomeação* e *tradução*, Tarski simplesmente constrói uma metalinguagem particular $\mathcal{L}'$, para estrutura de partida $\langle \mathcal{L}, \mathcal{A}, \mathcal{C} \rangle$, na qual algumas demandas relativamente vagas e intuitivas são claramente satisfeitas, demandas cuja satisfação pode ser apenas reconhecida na metaciência $\mathcal{C}'$ associada a $\mathcal{L}'$.

Dois exemplos relativamente simples (as estruturas de partida, $\langle \mathcal{L}, \mathcal{A}, \mathcal{C} \rangle$, e chegada, $\langle \mathcal{L}', \mathcal{C}' \rangle$, $V(x)$, $n(x)$, $t(x)$), são, supostamente, dadas:

Teorema II. Seja ‘ $\wedge$ ’ um conectivo binário de $\mathcal{L}$, então, para quaisquer sentenças α e β de $\mathcal{L}$, existe um termo ϵ de $\mathcal{L}'$ tal que $\vdash_{\mathcal{C}'} \epsilon(n(\alpha), n(\beta)) = n((\alpha \wedge \beta))$.

Teorema III. Para qualquer sentença α de $\mathcal{L}$, $\vdash_{\mathcal{C}} \alpha$ se e somente se $\vdash_{\mathcal{C}'} t(\alpha)$.

2.5 Metaciência tarskiana

No caso da “teoria” tarskiana da verdade, como ficou, pelo menos, subentendido, a metaciência e a metalinguagem tarskiana, os nomes descritivos-estruturais e a tradução tarskiana fazem sentido apenas conjuntamente e, conjuntamente, res-peitam certas demandas intuitivas. Contudo, poderíamos argumentar, há algo de circular aqui. O bom funcionamento conjunto da estrutura estaria apoiado por considerações do tipo: $\mathcal{L}'$ é realmente (o que poderíamos pensar como) uma metalinguagem para $\mathcal{L}$ (uma vez que, entre outras coisas, a capacidade expressiva de $\mathcal{L}'$ é superior a de $\mathcal{L}$ e que $\mathcal{L}'$ incorpora termos descritivos-estruturais como nomes de expressões de $\mathcal{L}$); $\mathcal{C}'$ é realmente (o que poderíamos pensar como) uma metaciência associada a $\mathcal{C}$ e $\mathcal{L}$; t é realmente uma tradução das fórmulas de $\mathcal{L}$ em $\mathcal{L}'$, e assim por diante. Entretanto, que algo seja uma metalinguagem e etc. depende da relação complexa entre os elementos das estruturas de partida e chegada.

Mais especificamente, embora a metaciência tarskiana seja definida na metalinguagem que lhe é associada, a metalinguagem é construída em vista de certos resultados metacientíficos específicos.

Na metaciência $\mathcal{C}'$ deverá ser possível estabelecer:

- (a) A sintaxe das expressões de $\mathcal{L}$ módulo nomes descritivos-estruturais [cf. Teorema II acima].
- (b) O comportamento das sequências de satisfação associadas às cláusulas quantificacionais tarskianas⁸.

A metaciência e, portanto, a metalinguagem tarskianas, os nomes estruturais descritivos e a tradução tarskiana são construídos em vista do estabelecimento do predicado tarskiano de verdade.

⁸Nesse caso, as relações entre tradução e nomes estruturais-descritivos é essencial.

Não se trata, portanto, da axiomatização de um domínio de significação constituído independentemente, como acontece com teorias formais usuais (aritmética ou análise formal), nem, portanto, da redução definicional de algum conceito nesse domínio, como no caso da divisibilidade no contexto da aritmética ou dos conceitos da análise e da aritmética no contexto da teoria dos conjuntos.⁹

Em outras palavras, não se trata *prima facie* da axiomatização do predicado independente de verdade e nem da axiomatização do predicado independente de verdade por meio definicional, mas da constituição do “predicado tarskiano de verdade” ou dos “predicados tarskianos de verdade” (se se quer enfatizar que cada estrutura de partida produz um predicado diferente).

2.6 Conclusão

O fato é que a “teoria” de Tarski é referida como definicional, entretanto, ela não pode ser definicional no sentido em que a aritmética formal de Peano é uma teoria definicional da divisibilidade ou ZF é uma teoria definicional dos conceitos da análise.

É claro que os predicados tarskianos de verdade são, ao menos, potencialmente definições e que um predicado de Tarski não é um predicado qualquer, em-bora não se refira, como tentei argumentar, a um domínio de “significação” já constituído e, *a fortiori*, ao predicado de verdade de uma teoria informal da verdade. Um predicado tarskiano de verdade é ele próprio uma construção. É verdade que essa construção responde a certas demandas formais e materiais (a Convenção T é tão somente uma delas), entretanto, apesar das demandas, o predicado tarskiano de verdade não é uma versão formal do predicado de verdade da linguagem natural, mas um predicado que supostamente poderia induzir, no discurso informal ele mesmo, um novo “conceito”: a *verdade tarskiana* e, como tal, poderia ser julgado por critérios como a utilidade, etc..

Tarski (1956: 154-165) afirma que a incompatibilidade entre as leis da lógica clássica e a universalidade característica das linguagens naturais coloca dificuldades insuperáveis à construção de uma definição de verdade adequada às linguagens naturais e que deveríamos, portanto, nos restringir, nas palavras do autor, “*inteiramente às linguagens formalizadas*” (ibidem: 165). Contudo, não se trata aqui de uma mesma noção de verdade, embora, de alguma forma, a verdade tarskiana se inspire, por meio da adequação a demandas implícitas e explícitas, em certas concepções de como um predicado de verdade deveria se comportar.

Scott Soames (1999: 56) afirma sobre a recusa de Tarski de atacar as “dificuldades insuperáveis” referidas anteriormente que não importa quão plausíveis sejam as assunções tarskianas, “[...] *elas são conjuntamente incoerentes e, portanto, inaceitáveis. De modo que a tarefa de encontrar princípios mais acurados e aceitáveis* [que aqueles pressupostos no argumento de Tarski] *permanece*”. Contudo, é exatamente isso que Tarski faz ... para as linguagens formalizadas. Como é sugerido por Wolfgang Künne (2003: 176-177), Tarski precisa de um conceito de verdade que satisfaça “*as necessidades da metodologia das ciências dedutivas*” (ou seja, o estudo dos conceitos formais de validade, consequência e completude) e que não leve ao paradoxo. Uma das possíveis respostas a essas necessidades mais gerais é (pelo menos, segundo o próprio Tarski) a verdade tarskiana.

Finalmente, note-se que, nesse contexto, boa parte da discussão lógico-filosófica sobre a “verdade” é uma discussão sobre as demandas formais e materiais que um *novo* predicado (tarskiano,

⁹Há uma sutileza aqui: dadas codificações gödelianas da linguagem formal em causa e contrapartidas conjuntistas dos “significados” das constantes dessa linguagem, podemos, de fato, estabelecer contrapartidas conjuntistas de vários “predicados de verdade”; em todo caso, se simplesmente emularmos a construção da estrutura tarskiana de chegada, esse predicado de verdade será apenas uma versão conjuntista do predicado *tarskiano* de verdade.

kripkeano, etc.) de verdade, supostamente mais adequado do que nosso predicado intuitivo e independente, estaria sujeito. Além disso, nada, à primeira vista, impediria, embora constituída definicionalmente de modo *sui generis* por Tarski (como nossa discussão enfatiza), a formalização direta e axiomática da noção tarskiana de verdade como conceito primitivo.

Em suma, a definição tarskiana do predicado de verdade pode ser pensada como estabelecimento do “conceito de verdade” ao qual essa definição se aplica.

Bibliografia

- Davidson, D. (2006). “Truth and meaning”, em *The Essential Davidson*, Oxford University Press, Oxford.
- Kirkham, R. (2001). *Theories of Truth: a Critical Introduction*. MIT press, Cambridge, Massachusetts.
- Künne, W. (2003). *Conceptions of Truth*. Clarendon Press, Oxford.
- Soames, S. (1999). *Understanding Truth*. Oxford University Press, Oxford.
- Tarski, A. (2006). *A Concepção Semântica da Verdade*. Orgs. C. A. Mortari, L. H. de Araújo Dutra. Editora da Unesp, São Paulo.
- Tarski, A. (1956). *Logic, Semantics and Metamathematics*. Translated by J. H. Woodger. Clarendon Press, Oxford.
- Tarski, A. (1935). “Der Wahrheitsbegriff in den formalisierten Sprachen”. *Studia Philosophica*, Vol. I, 261-405.

What is Logical about the Logical Interpretation of Probability?

Parzhad Torfehnezhad

Département de Philosophie
Université de Montréal
Pavillon 2910, boul. Édouard-Montpetit
Montréal, QC, H3C 3J7, Canada
parzhad.torfeh-nezhad@umontreal.ca

Abstract

My goal, in this paper, is to critically assess the categorization of “interpretations of probability” as it appears in the literature. In some sources only Carnap’s treatment of probability is understood to be the best example of “logical” probability. This is surprisingly narrow and I will here suggest otherwise. In fact, I believe that certain forms of Bayesianism should also be included in the logical camp.

Introduction

In the development of scientific theories, many pre-theoretical concepts, such as weight, volume, density, etc., acquire a precise theoretical meaning together with a systematic numerical assignment. The assignment normally takes place by defining a functor¹ that relates the domain of the physical properties such as weight to that of numbers. It is by the means of this apparatus that we make sense of the sentences such as “the weight of X is n ”. Clearly, such systems presuppose the basic logical and mathematical syntax of the relevant number theory they employ. In addition to these basic elements, certain primitive terms (e.g., m for mass, a for acceleration, etc.) and axioms (e.g., laws of classical mechanics) are added and constitute the theory. Semantical rules of designation are then laid down allowing us to interpret the primitive term m as mass, a as acceleration, and so on. In short, we may regard the full interpretations of pre-theoretic concepts as being constituted in the following two separate rudiments: Full Interpretation = (axioms & primitive terms) + (descriptive & pure semantical rules of designation)². There are some important points that one should notice here. First, it is well known that the symbolic calculus is to be developed independently of any interpretation of it. Second, both alternative set of axioms and alternative set of designation rules could lead to alternative interpretations.

¹The term “functor” is used in a Carnapian sense (Carnap, 1942: 17) not as it is commonly used in category theory. To see other examples of the use of axiomatic methods in science, see the section “Physical Calculi and Their Interpretations” in (Carnap, 1939: 56).

²Despite using Carnap’s terminology for the sake of coherence in this paper, I should clarify that the use of axiomatic method and the relationship between the primitive terms, axioms, and their interpretations as described in this paragraph, are not limited to Carnap’s. It is a commonly accepted concept in today’s mathematics and in mathematical textbooks (Lee, 2013; Novikov, 2001).

These considerations apply *mutatis mutandis* to the concept of probability. Indeed, what does it mean to say: “the probability of X is p ”? Probability is a vague concept we use in ordinary language that needs to be explicated³.

My goal, in this paper, is to critically assess the categorization of “interpretations of probability” as it appears in the literature, as, for instance, in the entry of the Stanford Encyclopedia of Philosophy (Hájek, 2012)⁴. In section 3 of “Interpretations of Probability” in the Stanford Encyclopedia of Philosophy, the author categorizes interpretations of probability as follows: 3.1 Classical probability, 3.2 Logical probability, 3.3 Subjective probability, 3.4 Frequency Interpretations, 3.5 Propensity Interpretations, and 3.6 Best-System Interpretations. According to the author, “early proponents of logical probability include (Johnson, 1921), (Keynes, 1921), and (Jeffreys, 1939). However, by far the most systematic study of logical probability was by Carnap”. I should mention that my objection to this categorization is not merely terminological, but, as we will see, it is directed to a seemingly-overlooked and important issue in identifying interpretations in axiomatic systems in general, being the possibility of having different interpretations on the basis of different sets of syntactic rules as opposed to having different interpretations due to different sets of semantic rules. Interestingly enough, Carnap himself may have pointed out this very issue, and if true, his philosophical position on probability should be re-evaluated, regardless of his proposal for inductive logic. It should be underlined that there is also a general point to be made about the relationship between interpretations, syntax and semantics in the axiomatic systems that is not limited to Carnap, and is a well-known fact in today’s mathematics.

As I have just said, only Carnap’s treatment of probability is considered to be the prime example of “logical” probability. This is surprisingly narrow. Can it be that only Carnap well defended a logical interpretation of probability? I will here suggest otherwise. In fact, I believe that certain forms of Bayesianism should also be included in the logical camp. I will also suggest another basis for classifying interpretations of probability according to which the subjective interpretation can also be understood as being “logical”. I should clarify here that the point of this paper is not to give a final answer to what probability is; it is neither the point to compare the advantages and disadvantages of different interpretations of probability. The main issue is to establish a framework, consistent with the present axiomatic methods, according to which one could clearly situate and group different interpretations of probability, and thereby have a better understanding of what logical probability is and how we should employ the term.

I will proceed as follows: I will start by considering probability theory as an axiomatized theory. Then, two possible and different sets of axioms in the theory will be given: one with the logical relation of consequence included, and one without. I will argue that at the very basic level, even before reaching the full interpretation of what probability is (that is to say considering semantic rules), there are two basic possibilities for final interpretation of probability. I will then conclude that this basic split should be the primary basis for categorizing interpretations of probability.

³The whole process of replacing vague concepts such as one-place predicates like “heavy” in “this thing is heavy” or a two-place predicate “heavier” in “this thing is heavier than that” by a functor that allows us to say, in a more precise way, “the weight of this thing is 7kg” is what Carnap calls an explication. In the given example, the vague concept “heavy” is called an explicandum and the clearer concept “weight (in kilogram)” is called an explicatum. At the axiomatic level, since there is no factual content involved, all that matters is the mathematical properties (such as consistency, completeness, independence, etc.) of the calculus upon which the number assignment is delivered. See (Carnap, 1947a: 7).

⁴This is but one important and typical example. For other examples see (Hájek & Joyce, 2008; Lyon, 2009; Sarkar & Pfeifer, 2006).

I will then proceed to Carnap's general analysis of probability. In this section, I will draw reader's attention to the two layers of Carnap's analysis of probability. The first layer concerns a linguistic approach to the term "probability" in ordinary language⁵. This first layer of the discussion acknowledges the existence of two fundamentally different concepts (two explicanda) of probability, one logical and one factual. I believe that this fundamental distinction speaks to the same point I am trying to make here concerning the two possible arrangements of axioms in probability theory. The second layer of the discussion, however, involves Carnap's own proposal for covering the logical concept. Although I will give a summary of Carnap's proposal for the logical aspect of probability (inductive logic), I will not defend, nor oppose, Carnap's position in this regard here⁶.

Probability Theory: A Brief Historical Background

In the mathematical sense, a probability space is a measure space. Measurements, in general, consist of assigning numerical values to different possibilities of measurable objects (in a continuous way). The first axiomatization of the theory was proposed by Andrey Kolmogorov in 1933 (Hájek, 2012), according to which probability P is considered to be a function from the set F to the set of real numbers satisfying the following conditions:

1. (Non-negativity): $P(A) \geq 0$, for all $A \in F$.
2. (Normalization): $P(\Omega) = 1$.
3. (Finite additivity): $P(A \cup B) = P(A) + P(B)$ for all $A, B \in F$ such that $A \cap B = \emptyset$.

Where Ω is a non-empty set (universal set), and F is an algebra⁷, closed under complementation and union, which is a subset of the power set of Ω . Now, let us take a look at a very brief background of some of the interpretations of probability. According to Bernoulli-Laplace's classical interpretation⁸, in a random process, if N is the number of equally-likely and mutually-exclusive outcomes, and N_A is the number of outcomes in which the desired event A occurs, then the probability of A would be calculated as the ratio of N_A to N .

$$\frac{P(A) = N_A}{N}$$

Frequentists, instead, focus on finding probability for actual events disregarding calculating probabilities prior to trial experiments. If a frequentist wants to know the probability of coming up a head in the process of coin-tossing, for example, she would actually toss the coin for a number of times and record the results in order to establish what is called a "reference class"; the greater the population of the reference class, the more accurate the probability. Then she would compare the limit of the ratio of frequency of the desired event in the reference class population to the total population of the reference class (when the population approaches to infinity) in order to calculate the probability of the desired event occurring. Therefore, if n_A

⁵This layer of discussion is directly related to Carnap's general analysis of linguistic frameworks in "the Foundations of Logic and Mathematics" (Carnap, 1939).

⁶This discussion would be a lengthy one in which one has to discuss more detailed subject matters such as whether or not the value of λ , and hence the confirmation function is regarded as completely arbitrary.

⁷An algebra, here, broadly speaking, is considered as a set along with some operations satisfying certain conditions. An n -ary operation on Ω is a *function* that takes n elements of Ω and gives a single element of Ω .

⁸This interpretation is famously circular for the use of the adjective "equally-likely" in the definition of probability, which itself is a measurement of likelihood. The problem would not be solved, even by appealing to the "principle of indifference": whenever there is no evidence favoring one possibility over another, they have the same probability (Hájek, 2012). It is worth mentioning that Carnap strongly rejects both the classical interpretation and any appeal to this principle: "We regarded the classical conception of probability, represented chiefly by Jacob Bernoulli and Laplace, as definitely refuted by the criticism of the frequentists. The classical conception was essentially based on the principle of insufficient reason or indifference according [...]", Carnap in (Schilpp, 1963: 70).

is the number of occurrences of A in the total number of n_t trial, then the probability of A would be $P(A)$ in the following formula where p is a real number:

$$P(A) = \lim_{n_t \rightarrow \infty} \frac{n_A}{n_t} = p$$

As practical as this interpretation might be, there are problems (both practical and theoretical) associated with it, *if* we limit ourselves just to this interpretation of probability. For example, in the case of the infinite ways of establishing a reference class, what would justify our choice of a reference class? In the case of coin tossing, is the reference class the class of infinite numbers of tossing the same coin, or it is infinite tossing of different coins of the same type (each only once)? Therefore, what would be the reference class of singular events? There are other problems associated with this interpretation that are discussed at length in the literature (Hájek, 2009).

On the other hand, total reliance on experimentation in order to capture the whole concept of probability may raise other problems, too. For instance, the obviousness of the sentence “when there is a 10% chance for the occurrence of A , then 90% of the time A would not occur” does not rely on any sort of empirical experiment. One can clearly see that there are two different applications of probability here; in one application, probability can be assigned to events synthetically via observation, whereas, in the other application, the assignment of probability follows some well-defined analytic structures⁹. We will have a more detailed discussion on this topic in the subsequent sections.

1 Logicality and Conditional Probability

One of the philosophical issues, with regard to conditional probability, is the issue of possibility versus probability. One may say “possibility” is a stronger concept than probability in the sense that the probability of impossible events might be considered as zero, but zero probability does not imply impossibility. According to Fitelson et al., Kolmogorov himself was certain that “probability 0 does not imply impossible” (Fitelson et al., 2006). For Kolmogorov, as we saw above, conditional probability $P(A|B)$ can be derived from unconditional probability (as represented above in the form of a one-place function $P(_)$).

It is obvious that, according to this formulation the conditional probability is undefined if either $P(B) = 0$ or any unconditional probabilities are undefined.

This situation, in which “zero-probability” is undefined, poses a problem for the theories in which probability is considered as a one-place function because, as Fitelson (Ibid) points out, “in uncountable spaces there can be genuine, non-trivial events whose probabilities are undefined (so-called ‘non-measurable sets’), and others whose probabilities are 0”.

Let’s take a short detour in order to clarify the zero-probability issue and why countable additivity becomes an important issue. Suppose we define the following:

- Sample space $\equiv \Omega$
- Sample point $\equiv$ possible outcome $\equiv \omega \in \Omega$
- Event $\equiv$ (a subset of Ω) $\equiv E$

Therefore, if $\Omega = [0, 1]$ then an event E is a subinterval $[a, b]$ of $[0, 1]$. Now, the property that the measuring function $P(x)$ ought to satisfy with respect to the subintervals is that if $[a, b] \subseteq [0, 1]$, then $P([a, b]) = b - a$. Accordingly, for a single possible outcome ω we have

$$\forall \omega \in \Omega, P(\{\omega\}) = P([\omega, \omega]) = \omega - \omega = 0$$

⁹The terms “analytic” and “synthetic”, here, are meant to be in the Carnapian sense.

This means that every possible outcome is a zero-probability event. This conclusion seems contradictory because if $\Omega = \cup\{\omega\}$, then $P(\Omega) = P(\cup\{\omega\}) = \sum P(\omega) = 0$, and we know $P(\Omega)$ is supposed to be 1. Here is where the axiom of countable additivity becomes important because the axiom only applies to countable objects and Ω is not countable.

Nevertheless, the zero-probability situation is a mathematically non-trivial question and it is meaningful in an infinite context; and if we are to assume probability as a one-place function, we have to face this situation in one way or another.

An alternative way of axiomatizing the notion of probability is to consider it not as a one-place function but as a two-place function, $P(_, _)$, and take this conditional probability as a primitive term (as Carnap does; see below). It has been shown (Stalnaker, 1970) that, in this case, it is possible to derive the unconditional probability of A as $P(A, T)$, where T is a logical truth (a L -true proposition, a tautology). According to Fitelson (Fitelson et al., 2006), various axiomatizations in which conditional probability is taken to be primitive have been defended in the literature¹⁰; and there are also canonical models with proven completeness and soundness in which probability is taken to be a two-place function (Lepage & Morgan, 2003).

A group of interpretations, also known as Bayesianism (Hájek, 2012), adopts this alternative axiomatization. Subjective interpretation is one of the examples in which conditional probability is primitively taken to be a two-place function. In the subjective interpretation of probability, it is the amount of the agent's knowledge (who is facing an uncertain situation) that plays a crucial role in the probability assignments. In other words, the source of uncertainty is considered to be the epistemic state of the agents. Therefore, the assignment of probability ought to be relativized to the amount of the knowledge the agent has. Then, in fact, what probability represents is the degree of the agent's belief. And, probability theory can be regarded as a guide to rational beliefs in every particular situation. Consequently, subjectivists heavily lean toward the concept of conditional probability (Joyce, 2008). Bayes' theorem¹¹ is central to subjectivists' theories of confirmation "both because it simplifies the calculation of conditional probabilities and because it clarifies significant features of subjectivist position" (Joyce, 2008).

Nowadays, Bayesian interpretation is regarded to be an extension of propositional logic¹². Hájek shows (Hájek, 2008) that even at the axiomatic level we may see the reliance of probability theory on deductive logic¹³. At the axiomatic level, we could, as well, attach probabilities to members of a collection S of *sentences* of a formal language, closed under (countable) truth-functional combinations, and consider counterparts of the above-mentioned Kolmogorov's axioms as the following:

- I. $P(A) \geq 0$ for all $A \in S$.
- II. If T is a logical truth (in classical logic), then $P(T) = 1$.

¹⁰There are also examples in which both conditional and unconditional probabilities are simultaneously axiomatized (Goossens, 1979) but the system is considered as an extension of Kolmogorov probability theory, and has the usual definitions linking conditional and unconditional probabilities as theorems.

¹¹According to the Bayes' theorem conditional probability can be considered as follows:

$$\frac{P(A, B) = P(B, A)P(A)}{P(B)}$$

It can be proven that Bayes' theorem is equivalent to the following formulation (Fitelson et al., 2006) (where $P(A^c)$ is the probability of the events in which A would not occur), which would not face the zero-probability problem:

$$\frac{P(A|B) = P(B, A)P(A)}{P(B, A)P(A) + P(B, A^c)P(A^c)}$$

¹²See (Leitgeb, 2014), (Pearl, 1991), (Cowell, 1999), (Jensen, 2001), (Kersting, 2007).

¹³For an extensive and detailed discussion on the alternative set of axioms including the conditional probability function as primitive see §4.2.1 (Leitgeb, 2015).

- III. $P(A \vee B) = P(A) + P(B)$ for all $A \in S$ and $B \in S$ such that A and B are logically incompatible.

In this case, since the notions of “logical truth”, “logical incompatibility” and “implication” are well defined in deductive logic, one can consider probability theory as a dependent theory on classical logic.

If we consider the evidence e and the hypothesis h as the sentential arguments of the probability function $P(e, h)$, the relativized Kolmogorov’s set of axioms would look like the following:

Kolmogorov’s Axioms	Relativized Kolmogorov Axioms
I. $P(E) \geq 0$, for all $E \in F$. II. $P(\Omega) = 1$. III. $P(E_1 \cup E_2) = P(E_1) + P(E_2)$ for all $E_1, E_2 \in F$ if $E_1 \cap E_2 = \emptyset$.	I. $P(e) \geq 0$ for all $e \in Z$. II. If T is a logical truth, then $P(T) = 1$. III. $P(h \vee e) = P(e) + P(h)$ for all $e \in Z$ and $h \in Z$ such that e and h are logically incompatible.

Table 1: Relativized Kolmogorov’s Axioms to Classical Logic.

Van Fraassen (Van Fraassen, 1995), for example, offers a probability system in which a two-place function is taken as primitive and the basic axioms for conditional probability were adopted rather than conditional probability is being defined in the usual way from monadic probability. Costa and Parikh (Costa & Parikh, 2005) offer an extended version of van Fraassen’s model and use it to model Bayesian update in a non-monotonic way. As quoted by Costa and Parikh (Ibid), van Fraassen studies two place probability functions $P(.|.)$ defined on a σ -field F over some set U with the following requirements:

- (I) For any fixed A , the function $P(X|A)$ as a function of X is either a (countably additive) probability measure, or has constant value 1.
- (II) $P(B \cap C|A) = P(B|A) P(C|B \cap A)$ for all A, B, C in F . If $C \subseteq B \subseteq A$, then (II) can be simplified as
- (II’) $P(C|A) = P(C|B) P(B|A)$.

As we see, mathematically speaking, as far as the conditionality is concerned, subjective interpretation can be well regarded as an extension of propositional logic; because, probability, right from scratch, is taken to be a conditional function. According to the above explanations, one can see that what can be deemed as “logical” in all of the possible interpretations on the basis of taking a two-place function as a primitive function, is the primitive inclusion of the logical consequence relationship in the axioms (i.e., conditional probability). In fact, Leitgeb (Leitgeb, 2014), based on similar argumentation, proposes a new category of logical systems called “probabilistic logic”. Maher (Maher, 2006, 2010), also talks about the legitimacy of using “inductive probability” as an umbrella term to cover the logical aspect of probability as opposed to what he calls “physical probability”.

So, in terms of axiomatization, we have basically two options with regard to conditional probability. We are either to take the absolute probability (a one-place function) as a primitive term and derive conditional probability from it as a theorem, or to take conditional probability primitively (as a two-place function) and derive absolute probability from it by adding some other definitions. Mathematicians have tried both of these options. It is fairly obvious that the choices are different and not completely equivalent, not only mathematically but also

philosophically. In one case, probability is taken to be an intrinsically conditional concept (requiring two arguments for a two-place function) and it is a primitive part of axiomatization, while, in the other case, conditional probability is not a part of axiomatization, and is derivable from an unconditional concept of probability (a one-place function). Now, one may easily see that regardless of the possible alternative semantic set-ups for each possibility, just at the axiomatic level, we can be sure that there exist at least two different (incompatible) possibilities for interpreting probability without knowing what the interpretations actually are going to be. This fact, as we will see in the subsequent section, might be related to what Carnap calls “two explicanda of probability”.

It should be fairly clear by now that if we are to categorize interpretations of probability at the end, the primary division should be on the basis of inclusion or exclusion of a two-place function in the axiom set as our measuring function: hence, the division would be between logical (in the sense of primitive inclusion of consequence relationship in the axiom set) versus non-logical ¹⁴ (in the sense of primitive exclusion of the conditional relation) probability. Once again, I would like to emphasize that this division is totally independent of any full interpretation of the concept; it is just the result of some mathematical manipulations regardless of the content. Therefore, it should be well regarded as a fundamental split rather than a semantical one. In the subsequent section, we will see some evidence, which strongly suggests that Carnap was already well aware and assertive of this very primary division with regard to the concept of probability, and perhaps it is according to this fundamental split that he finds the subjectivists’-frequentists’ debate futile.

In the literature, Carnap’s position is categorized as being a “logical interpretation” of probability, meaning that it is neither Bayesian, nor frequentist (Fitelson et al., 2006; Hájek, 2012; Lyon, 2009; Sarkar & Pfeifer, 2006). As we discussed above, being logical (or not) is more a theoretical matter rather than a matter of descriptive semantics, hence, considering this fact, certain forms of Bayesianism can also be equally considered as “logical” for their theoretical positions with regard to conditional probability. Therefore, the common categorization of interpretations of probability in the present literature does not seem to be completely correct¹⁵. It seems to me that the underlying assumption behind this way of categorizing the various interpretations of probability (contrary to Carnap’s analysis) is that probability is a unique and single explicandum that needs to be explicated. And I assume it is probably due to this assumption that the list of interpretations of probability is given generally in the literature, e.g., (Lyon, 2009; Sarkar & Pfeifer, 2006), and in particular in the above-mentioned entry in Stanford Encyclopedia of Philosophy. Thus, we read in the very first lines of the first paragraph.

¹⁴Instead of non-logical, one may use the term “frequential” or “physical”.

¹⁵I should mention here that Hájek himself is against using the term “interpretation” in the sense we normally use with respect to formal systems on two grounds: First, there are other interpretations of probability that are not exactly based on Kolmogorov’s Axioms yet are categorized under it (perhaps under the logical setting we talked about). Secondly, Kolmogorov’s Axioms can be equally employed for measuring spaces other than probability.

Normally, we speak of interpreting a formal system, that is, attaching familiar meanings to the primitive terms in its axioms and theorems, usually with an eye to turning them into true statements about some subject of interest. However, there is no single formal system that is ‘probability’, but rather a host of such systems. To be sure, Kolmogorov’s axiomatization [...] has achieved the status of orthodoxy [...]. Nevertheless, several of the leading ‘interpretations of probability’ fail to satisfy all of Kolmogorov’s axioms, yet they have not lost their title for that. Moreover, various other quantities that have nothing to do with probability do satisfy Kolmogorov’s axioms, and thus are interpretations of it in a strict sense: normalized mass, length, area, volume, and other quantities that fall under the scope of measure theory, the abstract mathematical theory that generalizes such quantities. Nobody seriously considers these to be ‘interpretations of probability’, however, because they do not play the right role in our conceptual apparatus. (Hájek, 2012) But there is no need to modify the meaning of “interpretation” if we accept that there are fundamentally two different sets of axioms for measuring probability and they are both related to the concept of measurable space.

‘Interpreting probability’ is a commonly used but misleading characterization of a worthy enterprise. The so-called ‘interpretations of probability’ would be better called ‘analyses of various concepts of probability’, and ‘interpreting probability’ is the task of providing such analyses. Or perhaps better still, if our goal is to transform inexact concepts of probability familiar to ordinary folk into exact ones suitable for philosophical and scientific theorizing, then the task may be one of ‘explication’ in the sense of Carnap.’ (Hájek, 2012: 1)

Probability in Carnap’s Philosophy

After establishing the existence of two fundamentally different settings for interpreting probability, let us look at Carnap’s treatment of probability. One should be aware that there are two levels in Carnap’s discussion on probability. At the first level Carnap takes a linguistic approach and considers the term “probability” as an explicandum. This approach is, of course, closely correlated to his general approach in analyzing philosophical problems with respect to the concept of linguistic framework and his theory of meaning, which would require more elaboration (see (my own paper) for details). It is here, at this level, that I believe he wants to authenticate what we have established so far, namely, the existence of two settings for interpreting one concept, probability, or, using his own vocabulary, the existence of two explicanda for one term “probability”. Once he establishes that there is a logical setting (logical explicandum) for the term, *then*, we go to the second level of his analysis where he proposes his method for explicating this explicandum via what he calls “inductive logic”.

Carnap’s philosophy, as we will see, suggests that some concept cannot be explicated as one single explicandum and, probability is one of them. For Carnap both logical and factual explicanda of probability need to be explicated.

Carnap believes that there are two fundamentally different explicanda of probability, each needing to receive explications separately, and hence, each amount to a different kind of confirmation. This, in my view, is the first level in Carnap’s philosophy of probability, as I mentioned above. But, if we accept this premise, then we ought to accept that no sole theory of probability can deliver the whole meaning of the term: to capture the whole meaning of probability one needs at least two different theories.

‘The various theories of probability are attempts at an explication of what is regarded as the prescientific concept of probability. In fact, however, there are two fundamentally different concepts for which the term ‘probability’ is in general use. [...] (i) Probability₁ is the degree of confirmation of a hypothesis *h* with respect to an evidence statement *e*, (ii) Probability₂ is the relative frequency (in the long run) of one property of events or things with respect to another. A sentence about this concept is factual, empirical. Both concepts are important for science. Many authors who take one of the two concepts as explicandum are not aware of the importance or even of the existence of the other concept. This has led to futile controversy. (Carnap, 1962: 19)

It is well-known that for Carnap “There are two explicanda, both called ‘probability’: (1) logical or inductive probability (probability₁), (2) statistical probability (probability₂)” (Carnap, 1973: 269). Carnap is quite clear in that “there is no one meaning of the term ‘probability’ which is applied with perfect consistency throughout his work by any of the classical authors” (Carnap, 1962: 50). He believes that in some cases, when we have large-enough reference classes, the relative frequency can be regarded as the representative of the ultimate relative fre-

quency, but this concept still ought to be regarded as the explicandum of the factual probability (probability₂).

‘If we take a sufficiently large unknown class K, then the relative frequency of M in K may be regarded as representing the relative frequency “in the long run”. But this is the explicandum of probability₂, the statistical concept of probability. (Carnap, 1962: 173)

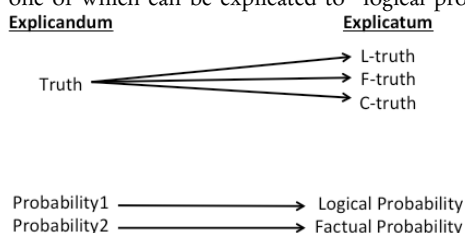
It is worth mentioning that Carnap makes a sharp distinction between “logical” and “factual truth”, and also between “pure” and “physical geometry”. Making those distinctions are not quite similar to the way he distinguishes between explicanda of probability¹⁶ (as “a concept mostly used by scientists and a concept suited for philosophical investigations”). But, sorting out concepts in a factual-theoretical continuum is characteristic of Carnap’s approach towards many philosophical problems. And it is tied to his overall position with regard to linguistic framework, abstraction and his whole theory of meaning. Making these distinctions is not only philosophically important for Carnap but also quite beneficial, because:

In this way we obtain also a clear distinction between questions about contingent facts and questions about meaning relations. This difference seems to me philosophically important; answering questions of the first kind is not part of the philosopher’s task, though he may be interested in analyzing them; but answers to questions of the second kind lie often within the field of philosophy or applied logic. Carnap in (Schilpp, 1963: 63)

There are many places in which Carnap refers to frequentist interpretations as legitimate and acceptable interpretations of factual probability. The main point is that the job of dealing with this interpretation is not primarily the job of philosophers. It is true that, as a philosopher, Carnap tends to work on theories of logical probability, but that does not mean that the concept of factual probability is *persona non grata* in his philosophy. On the contrary, for Carnap, factual probability is a legitimate concept, but responsible for only one part of the whole meaning of probability, which may well receive its own treatment via a frequentist interpretation in light of the empirical data.

On the other hand, probability₂ is an empirical concept; it is the relative frequency in the long run of one property with respect to another. The controversy between the so-called logical conception of probability, as represented e. g. by Keynes, and Jeffreys, and others, and the frequency conception, maintained e. g. by v. Mises and Reichenbach, seems to me futile. These two theories deal with two different probability concepts which are both of great importance for science. Therefore, the theories are not incompatible, but rather supplement each other. (Carnap, 1945: 591)

¹⁶The difference between explicating “probability” and “truth”, for instance, is that “truth” is considered as a *single*-explicandum (and Carnap explicates it into three different explicata F-truth, L-truth, and C-truth). “Probability”, in contrast, is considered as a concept that has essentially two explicanda (or a *double*-explicandum, if you wish), one of which can be explicated to “logical probability” and the other to “factual probability” in a parallel manner.



The fundamental difference is rather this 'probability2' designates an empirical function, viz., relative frequency, while 'probability1' designates a certain logical relation between sentences; these sentences, in turn, may or may not refer to frequencies. (Carnap, 1946: 72)

The statistical concept of probability is well known to ail those who apply in their scientific work the customary methods of mathematical statistics. [...] In the simplest cases, probability in this sense means the relative frequency [...] Thus the statistical concept of probability is not essentially different from other disposition concepts, which characterize the objective state of a thing by describing reactions to experimental conditions [...]. (Carnap, 1955: 1)

Thus, not only does Carnap have no objection to the frequentist theory, but also he even sees it as a necessary complementary part of the general theory of probability. However, frequentist theory alone is incapable in delivering the whole meaning of probability, and the same is true for logical theories of probability. Understandably, Carnap gives an extensive treatment on the logical aspect of probability, which is briefly presented in the next section, but one has to keep in mind that in this treatment Carnap by no means claims that this theory alone would deliver the whole meaning of probability or that this is *the* interpretation of probability, since there is none.

2 Inductive Logic

In this section, we move to the second level of discussion about probability in Carnap's philosophy. For Carnap, inductive logic is not understood in its traditional, Aristotelian sense, where one deals with inferences from particulars to universals, but as an extension of deductive logic that deals with uncertain propositions. Carnap's reasoning for taking a linguistic stance towards probability rests on the assumption that our knowledge about the facts eventually ought to be expressed in the form of propositions. If the propositions are considered to be certain then deductive logic is sufficient, otherwise one should employ an inductive logic to study their relationships. The difference is depicted in the following diagram:

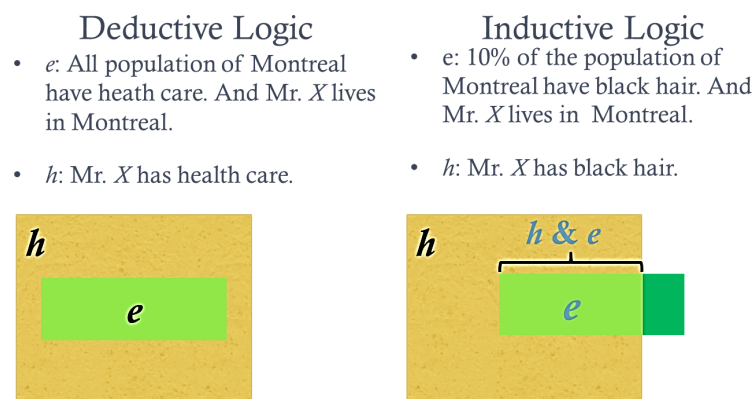


Figure 2.1: Deductive vs. Inductive Logic.

Subsequently, changes in the new evidence that may, in effect, change the previous probability assignment eventually will fall under one of the following events: subtraction or addition of (i) a sentence, (ii) a predicate, or (iii) an object name. We will see how the probability assignment will be subjected to change following each of these events.

For Carnap, "logical probability" is the central concept of inductive logic, in a similar way that "logical implication" is central to deductive logic.

Inductive logic is here understood as a theory based on a definition of the logical concept of probability, [...]. Its basic concept, the degree of confirmation, is in a certain sense a weak analogue of the concept of logical implication, the basic concept of deductive logic. (Carnap, 1947b: 133)

As we established above, subjective interpretations rest on deductive logic because of their dependency on logical concepts at the axiomatic level. The same is true for logical probability in Carnapian sense; “inductive logic is constructed from deductive logic by the adjunction of a definition of c [measuring function]. Hence inductive logic presupposes deductive logic” (Carnap, 1962: 192).

So, Carnap’s measuring function is called c -function and it is considered to be the explicatum for the vague concept of “logical probability” as its explicandum. C -function is considered to be intrinsically conditional and it is a function of two arguments: evidence e and hypothesis h . This concept, in turn, can play one of the following roles as:

- I. The *classificatory* concept (confirming evidence): $c(h, e) > b$.
- II. The *comparative* concept (higher confirmation): $c(h, e) > c(h', e')$, or $c(h, e) > c(h, e')$, or $c(h, e) > c(h', e)$.
- III. The *quantitative* concept (degree of confirmation): $c(h, e) = u$

Carnap then proposes the following five fundamental axioms that can be applied to any pairs of sentences e and h in a given language L (finite or infinite)(Carnap, 1962: 295).

Carnap’s Axioms for the c -function	Relativized Kolmogorov’s Axioms
I. Range of values: $0 \leq c(h, e) \leq 1$ II. L-implication: If $\vdash e \supset h$, then $c(h, e) = 1$ III. Special addition: If $h \not\vdash h'$ e is L-false, then $c(h \vee h', e) = c(h, e) + c(h', e)$ IV. General multiplication: $c(h \wedge h', e) = c(h, e) \times c(h', h \wedge e)$ V. L-equivalence: If $\vdash e \equiv e'$ and $\vdash h \equiv h'$, then $c(h, e) = c(h', e')$	I. $P(e) \geq 0$ for all $e \in Z$. II. If T is a logical truth, then $P(T) = 1$. III. $P(h \vee e) = P(e) + P(h)$ for all $e \in Z$ and $h \in Z$ such that e and h are logically incompatible.

Table 2: Carnap’s vs. Relativized Kolmogorov’s Axioms.

In order to measure the degree of confirmation Carnap introduces another measuring function called m -function that its regular version will be defined for a language L of N objects and containing state-description Z_i (Carnap, 1973), as follows:

m is a regular m -function for L N^{def}

(a) For every Z_i in L_N , $m(Z_i) > 0$

(b) $\sum_i m(Z_i) = 1$

(c) If j is L-false, $m(j) = 0$

(d) If j is not L-false, $m(j) = \sum_i m(Z_i)$ for all Z_i in the *range* of j

A state-description describes a (possible) state of affairs or model. Example: if a language system L contains symbols for only three objects a, b, c , and one monadic predicate F , then there exist three atomic sentences $i:Fa, j:Fb$, and $k:Fc$ along with their negations $\sim i, \sim j, \sim k$. A *state-description* Z_i is any conjunction of the mentioned atomic sentences that contains either an atomic sentence or its negation but not both. Example: $Z_1 = i \& \sim j \& k$.

Therefore, the disjunction of all state-descriptions ($\vee Z_i$) gives us a universal class, and the conjunction of them ($\wedge Z_i$) gives us the empty class (see Q-predicates below). The *range* of i in

L is the class of those Z in L in which i holds. So, the situation for state-descriptions in our example language of three objects (a, b, c) and one monadic predicate (F) can be summarized in the following table:

Z_i	State-description	m	Structure-description	Logical Weight	m^*
Z_1	$Fa \ \& \ Fb \ \& \ Fc$	1/8	Everything is F	1/4	1/4
Z_2	$\neg Fa \ \& \ Fb \ \& \ Fc$	1/8	Only two are F	1/4	1/12
Z_3	$Fa \ \& \ \neg Fb \ \& \ Fc$	1/8			1/12
Z_4	$Fa \ \& \ Fb \ \& \ \neg Fc$	1/8			1/12
Z_5	$\neg Fa \ \& \ \neg Fb \ \& \ Fc$	1/8	Only one F	1/4	1/12
Z_6	$\neg Fa \ \& \ Fb \ \& \ \neg Fc$	1/8			1/12
Z_7	$Fa \ \& \ \neg Fb \ \& \ \neg Fc$	1/8			1/12
Z_8	$\neg Fa \ \& \ \neg Fb \ \& \ \neg Fc$	1/8	Everything is not F	1/4	1/4

Table 3: State-Descriptions for L_3^1 .

Using c -function, Carnap then calculates the degree of confidence c (or c^* considering the logical weight) for the hypothesis h given the evidence e as follows:

$$c(h, e) = \frac{m(e \wedge h)}{m(e)} \quad \text{or} \quad c^*(h, e) = \frac{m^*(e \wedge h)}{m^*(e)}$$

Let's first examine the c -function (regular, i. e., all state-descriptions have equal m -value). Take the hypothesis $h = Fa$, thus the range is Z_1, Z_3, Z_4, Z_7 , and $m(h) = 1/2$. Let's the evidence be $e = Fb$, then its range is Z_1, Z_2, Z_4, Z_6 , and $m(e) = 1/2$. Then, the range of $h \wedge e$ is Z_1, Z_4 is $m(h \wedge e) = 1/4$. Hence, $c(h, e) = (1/4) / (1/2) = 1/2$ which is equal to $m(h)$ and thus it does not confirm it. On the other hand, if we consider all structure-descriptions we will calculate the c^* -function as follows:

$$m^*(h) = 1/4 + 1/12 + 1/12 + 1/12 = 1/2$$

$$m^*(e) = 1/2$$

$$m^*(h \wedge e) = 1/4 + 1/12 = 1/3$$

Hence $c^*(h/e) = m^*(h \wedge e) / m^*(e) = (1/3) / (1/2) = 2/3$. Then, by this function, the evidence confirms the hypothesis.

In our example, in which we have only one monadic predicate F , the only possible form of proposition is $P = F(x)$. This proposition may either have the property of $Q1$ (meaning P) or $Q2$ (meaning $\neg P$). Knowing these properties, known as Q -predicates or Q -properties, is essential in determining the logical width (w).

Predicate expression	Logical Nature	Q-Predicate	Width (w)	Relative width (w/k)
$P \wedge \neg P$	Empty		0	0
P	Factual	$Q1$	1	$1/2$
$P \vee \neg P$	Universal	$Q1 \vee Q2$	2	$1/2$

Table 4: Q -predicates for L_3^1 .

Structure-description Str_i is the disjunction of isomorphic Z_i in L that can be identified by Q -numbers, and expressed by a Q -predicate. In the case of having only three monadic predicates, for example, we may have three propositions: $P_1 = F(x)$, $P_2 = G(X)$ and $P_3 = H(x)$. Accordingly we will have the following Q -properties:

Str_i	Q-Predicates	Expression	Logical Nature	Width (w)	w/ κ
1		$Pi \wedge \neg Pi$	Empty	0	0
2	Q_1	$Pi \wedge Pj \wedge Pk$	Factual	1	1/8
3	$Q_1 \vee Q_2$	$Pi \wedge Pj$		2	1/4
4	$Q_1 \vee Q_2 \vee Q_3$	$Pi \wedge (Pj \vee Pk)$		3	3/8
5	$Q_1 \vee Q_2 \vee Q_3 \vee Q_4$	Pi		4	1/2
6	$Q_1 \vee Q_2 \vee Q_3 \vee Q_4 \vee Q_5$	$Pi \vee (Pj \vee Pk)$		5	5/8
7	$Q_1 \vee Q_2 \vee Q_3 \vee Q_4 \vee Q_5 \vee Q_6$	$Pi \vee Pj$		6	3/4
8	$Q_1 \vee Q_2 \vee Q_3 \vee Q_4 \vee Q_5 \vee Q_6 \vee Q_7$	$Pi \vee Pj \vee Pk$		7	7/8
9	$Q_1 \vee Q_2 \vee Q_3 \vee Q_4 \vee Q_5 \vee Q_6 \vee Q_7 \vee Q_8$	$Pi \vee \neg Pi$	Universal	8	1

Table 5: Q-properties for L with 3 primitive monadic predicates.

In general, for a language L_π^n where n is the number of individual constants, and π is the number of primitive monadic predicates the following holds:

- The number of atomic sentences is $\beta = \pi n$
- The number of Q-predicates is $\kappa = 2^\pi$
- The number of state-descriptions is $2^\beta = 2^{\pi n} = (2^\pi)^n = \kappa^n$

So far we were considering the finite situation. Now, if n and π equal to the cardinality of natural numbers $\aleph_0$, then κ forms a continuum, so does the number of state-descriptions.

In the case of continuum, Carnap introduces the function $\lambda(\kappa)$ that takes a non-negative real number to characterize the confirmation function $c\lambda(h, e)$. In the formula for a singular predictive inference¹⁷ to be given below the evidence statement, e_Q , says that s_1 individuals have property Q_1 , s_2 have Q_2 , ..., s_κ have Q_κ . The hypothesis h_i says that some individual not mentioned in the evidence has Q_i property ($s_1 + s_2 + \dots + s_\kappa = s$).

$$c(h_i, e_Q) = \frac{s_i + \left(\frac{\lambda(\kappa)}{\kappa}\right)}{s + \lambda(\kappa)}$$

Example: In a language that has two objects and one monadic predicate L_1^2 we have $n = 2$; $\pi = 1$; $\kappa = 2^\pi = 2$, and for the c^* -functions $\lambda^*(\kappa) = \kappa = 2$. If $e_Q = Fa$ and $h = Fb$ then $s_1 = 1$, $s_2 = 0$; $s = s_1 + s_2 = 1$. Accordingly, $c^*(h, e) = (1 + 1)/(1 + 2) = 2/3$. For this language, Q-predicates and state-descriptions (Z_i) are:

Predicate Expression	Logical Nature	Q-Predicate
$P \wedge \neg P$	Empty	
P	Factual	Q_1
$P \vee \neg P$	Universal	$Q_1 \vee Q_2$

Z_i
$Fa \wedge Fb$
$Fa \wedge \neg Fb$
$\neg Fa \wedge Fb$
$\neg Fa \wedge \neg Fb$

¹⁷According to Carnap, main kinds of inductive inferences for a language L_π^n are:

- Direct inference (from the population to a sample),
- Predictive inference (from one sample to another),
- Inference by analogy,
- Inverse inference (from a sample to the population),
- Universal inference (from a sample to a universal law).

There are two methods of considering λ . In the first method, λ is considered to be independent of κ . In that case, $c_\lambda(h_M, e_M) = (s_M + (w/\kappa)\lambda)/(s + \lambda)$. In the second method, λ can be considered as proportional to κ by a constant factor C , which means $\lambda(\kappa) = C\kappa$. In that case $c_\lambda(h_M, e_M) = (s_M + Cw)/(s + C\kappa)$. Regardless of the method, in both of these c -functions, Carnap recognizes the ratio of s_M/s as the *empirical factor* and that of w/κ as the *logical factor*.

In the exact same way as we saw above, Carnap derives unconditional probability (for a single statement j) from conditional probability given the evidence e is a tautological truth, and he calls it “null confirmation” c_0 (Carnap, 1973):

$$c_0(j) \stackrel{\text{def}}{=} c(j, t), \text{ where } t \text{ is a tautological evidence}$$

I need not go into further detail about Carnap’s interpretation to establish the point I want to make here. The point is that there are basic similarities between Carnap’s interpretation and certain forms of Bayesianism including subjective interpretation, namely: (a) taking probability as intrinsically a two-place relation between two claims, (b) the dependency on implication relationship, (c) deriving unconditional probability by assuming a logical tautology. And because of these similarities they should all be called “logical interpretations” of probability. In fact, we may say that any interpretation relying on any axiomatization of probability that satisfies the three mentioned conditions should be understood as being a logical interpretation of probability. As we saw, all such interpretations, at the axiomatic level (independent of any interpretation), are heavily dependent on the “implication relationship” which is understood to be a purely logical relationship, and this is the reason that allows us to put all these interpretations under the same umbrella. Whether or not subjective probability is a special case of Carnap’s logical analysis, or whether or not it can be derived from Carnap’s set of axioms, is irrelevant to the point that subjective interpretation, nevertheless, gives us a logical interpretation of probability. Therefore, if we agree that there is a basic split in the concept of probability axiomatically, then we have no choice but to accept Carnap’s first point about the essential duality of the concept of probability. And we will find, firstly, that the classifications of interpretations of probability we may see in sources such as the “Stanford Encyclopedia of Philosophy” need a substantial reevaluation, and secondly, that the controversy over whether subjective or frequentist interpretations give us the complete meaning of probability, is futile.

All in all, Carnap is clear in saying that a logical interpretation via conditional probability in the form of the “degree of confirmation”, in practice, would provide us the best estimation of a hypothesis, with respect to the given evidence that can be considered as a guide in life (Carnap, 1947c). And, that the assigned probability values to several hypotheses “can be interpreted as the estimate of the relative frequency of truth among them” (Carnap, 1962: 172).

3 The Proposal and Conclusion

To be clear, when we call an interpretation “non-logical” it does not mean that the interpretation in question is illogical or does not follow the rules of logic. As I mentioned in the introduction, any well-defined functor that assigns a number to a property necessarily assumes a logical structure. Therefore, in this sense all interpretations are logical. That being said, let us consider the separation criterion. According to our discussion thus far, one can say every interpretation of probability that satisfies the following three conditions should be understood as a logical interpretation. If an interpretation fails to satisfy these conditions, then it is a non-logical interpretation:

- a) Probability is a two-place relation between two claims.

- b) The axioms rely on the implication relation.
- c) Unconditional probability is derived by assuming a tautology.

Interesting consequences follow from this. For example, Bayesianism turns out to be logical in the same way as Carnap's inductive logic. One of the advantages of this way of classifying interpretations of probability is that it would allow one to primarily and fundamentally group interpretations of probability into two basic "logical" and "non-logical" categories. This would confirm the existence of the two explicanda for probability discussed above. Mathematically speaking, this means there exists at least two non-equivalent (not mutually exclusive) ways for delivering interpretations of probability. Linguistically speaking, this means that there are two different meanings associated with the word "probability". Thus, the concept of probability can be explicated in different ways. Accordingly, one cannot say there is one single interpretation that constitutes *the* interpretation of probability. This may finally put an end to some long lasting philosophical debates such as the subjectivists-frequentists debate over the true interpretation of probability. Of course there remain important questions concerning the nature of the relation between these meanings (e. g., whether they are complimentary or not) and how we are to accommodate these interpretations into our overall theory of probability. Satisfactory answers to these questions require some technical and theoretical adjustments.

The ramification of accepting two explicanda for probability will affect other fields of study as well. One major consequence concerns the confirmation of scientific (probabilistic) theories. One might be tempted to say that such a scientific theory may undergo two different kinds of evaluations. If probability is understood to be logical, then it means we are facing two claims, derived from the same theory, and we evaluate them with respect to each other. On the other hand, in the case of non-logical understanding, we may only consider one claim and evaluate that claim with respect to an established reference class (considering mathematical features such as stability, normalization, and the like). In the second case one considers rules and elements that are not necessarily embedded (or integrated) inside the scientific theory in question. Depending whether our interpretation is understood to be logical or not, the two evaluations are not the same. In the first case the result of the evaluation is responsive to the scientific theory alone. In the second case, the result is also responsive to an auxiliary mathematical theory. The latter is fundamentally different from the scientific theory. For example, in the case where we assign a 0.5 probability to a coin toss, one may evaluate this assignment by considering the description of all the circumstances in the actual process along with the two possible state-descriptions (as seen above in the section 4). On the other hand, one may establish a large enough reference class and observe that it would stabilize at 0.5. In the second case the stability in question is derived from a mathematical theory and disregards the empirical circumstances. That is to say we arrive at the stability on the assumption that this is not limited to the process of coin tossing (identified specifically by the corresponding state-descriptions), and that any set of data that behaves in this manner would be assigned the same ratio. Therefore, although the results of both evaluations are the same in this case, the evaluations themselves do not mean the same thing. There might be cases where only one type of evaluation is possible (e. g., cases in which one cannot possibly establish a reference class or a large enough reference class), and there are cases in which both types of evaluation are possible. If the evaluation of a probability statement is taken to be logical, then what is at stake is the logical consistency of the theory in light of the embedded implication relation (upon which rests the probability). When it is construed to be non-logical, only the validity of the reference class in question (or the sampling method, base lines, etc.) is at stake. It is obvious

that discussing confirmation is well beyond the scope of this paper; nevertheless, I just wanted to give the reader a sense of how this discussion might be affected given the validity of both logical and non-logical interpretations of probability.

I tried to establish, in this paper, that there are two different possibilities for an axiomatic setting with regard to probability at the *theoretical* level. The first is one in which the absolute probability is taken as primitive and conditional probability is derived as a theorem (the initial setting of Kolmogorov). The second is one in which it is the conditional probability that is taken as primitive and the absolute probability is derived (though not in the exact same way). Because of the inclusion of the logical consequence relationship in the later setting, one may meaningfully attribute the descriptor “logical” to all the interpretations based on this second setting. Hence the primary division in analysis of probability appears at the theoretical level, which is independent of any further interpretations of the systems. I believe that if we are to classify interpretations of probability, we ought to consider this fundamental split in our classification. And if we do so we will find Carnap’s inductive logic along with all forms of Bayesianism on the same side of the division.

I have also tried to show that there are two levels in Carnap’s philosophy of probability. At the first level, Carnap speaks of probability from a meaning-theoretic point of view, which bears strong resemblance to our discussion about the fundamental split in analyzing the concept of probability. In the second level, we deal with explicating *one* of the possible explicanda of probability. And, if we are to assign a name for this second part, we may call it “propositional interpretation of probability”¹⁸.

The exact nature of the relationship between the meanings of complete interpretations (based on each axiomatic setting) remains an open philosophical question. Nevertheless, Carnap’s answer to this question is that the meanings would be complimentary and there is no reason to adopt one at the expense of losing the other. If we agree with Carnap on this point, then, for the coherence of the theory, we ought to either adopt a principle that says the meanings of all such concepts (concepts with double explicanda) are complimentary (provided that each is based on only one of the axiomatic settings), or provide a criterion for identifying under which circumstances the meanings are complementary.

References

- Carnap, R. (1939). *Foundations of Logic and Mathematics* (Vol. 1, no. 3). Chicago: University of Chicago Press.
- Carnap, R. (1942). *Introduction to Semantics*. Cambridge, Mass.: Harvard University Press.
- Carnap, R. (1945). “On Inductive Logic”, *Philosophy of Science*, 12(2), 72–97.
- Carnap, R. (1946). “Remarks on Induction and Truth”, *Philosophy and Phenomenological Research*, 6(4), 590–602.
- Carnap, R. (1947a). *Meaning and Necessity: A Study in Semantics and Modal Logic*. University of Chicago Press, Chicago.
- Carnap, R. (1947b). “On the Application of Inductive Logic”, *Philosophy and Phenomenological Research*, 8(1), 133–148.
- Carnap, R. (1947c). Probability as a Guide in Life. *The Journal of Philosophy*, 44(6), 141–148.
- Carnap, R. (1955). “Statistical and Inductive Probability”, in: R. Carnap (Ed.), *Statistical and Inductive Probability & Inductive Logic and Science*. The Galois Institute of Mathematics and Art, Brooklyn, NY.

¹⁸This way, at least, one side of the debate between subjectivists and Carnapians might be reduced to a debate on the epistemic difference between propositions and beliefs.

- Carnap, R. (1962). *Logical Foundations of Probability* (Vol. 2d Chicago). University of Chicago Press, Chicago.
- Carnap, R. (1973). "Notes on Probability and Induction", *Synthese*, 25(3/4), 269–298.
- Costa, H. A. & Parikh, R. (2005). "Conditional Probability and Defeasible Inference", *Journal of Philosophical Logic*, 34(1), 97–119.
- Cowell, R. G., Dawid, P., Lauritzen, S.L., Spiegelhalter, D.J. (1999). *Probabilistic Networks and Expert Systems*. Springer, New York.
- Fitelson, B., Hájek, A. & Hall, N. (2006). "Probability", in: S. R. Jessica Pfeiffer, S. Sarkar (Ed.), *The Routledge Encyclopedia of Philosophy of Science*: Routledge, New York, 599–609.
- Goossens, W. (1979). "Alternative Axiomatizations of Elementary Probability Theory", *Notre Dame Journal of Formal Logic*, 20, 227–239.
- Hájek, A. (2008). "Probability - a Philosophical Overview", in: R. B. Gold, R. Simons (Eds.), *Current Issues in the Philosophy of Mathematics from the Perspective of Mathematicians*: Mathematical Association of America, Washington DC, 323–339.
- Hájek, A. (2009). "Fifteen Arguments against Hypothetical Frequentism", *Erkenntnis*, 70, 211–235.
- Hájek, A. (2012). "Interpretations of Probability", *The Stanford Encyclopedia of Philosophy* (Winter 2012 Edition), Edward N. Zalta (ed.), <http://plato.stanford.edu/archives/win2012/entries/probability-interpret/>
- Hájek, A. & Joyce, J. M. (2008). "Confirmation", in: S. Psillos & M. Curd (Eds.), *The Routledge Companion to the Philosophy of Science*: Routledge, New York.
- Jeffreys, H. (1939). *The Theory of Probability* (3 ed.). Oxford University Press, Oxford.
- Jensen, F. V. (2001). *Bayesian Networks and Decision Graphs*: Springer, New York.
- Johnson, W. E. (1921). *Logic*. Cambridge University Press, New York.
- Joyce, J. (2008). "Bayes' Theorem", *The Stanford Encyclopedia of Philosophy* (Fall 2008 Edition), Edward N. Zalta (ed.), <http://plato.stanford.edu/archives/fall2008/entries/bayes-theorem/>
- Kersting, K., De Raedt, L. (2007). "Bayesian Logic Programming: Theory and Tool", in: L. T. Getoor, B. Taskar (Ed.), *Introduction to Statistical Relational Learning*. MIT Press, Cambridge, MA, 291–318.
- Keynes, J. M. (1921). *A Treatise of Probability*. Macmillan and Co., Ltd, London.
- Lee, J. M. (2013). *Axiomatic Geometry*. Providence, Rhode Island: American Mathematical Society.
- Leitgeb, H. (2014). "Probability in Logic", in: A. Hájek & C. Hitchcock (Eds.), *The Oxford Handbook of Probability and Philosophy*: Oxford University Press, Oxford.
- Leitgeb, H. (2015). Preface to *The Stability of Belief: An Essay on Rationality and Coherence*. First draft published online https://www.academia.edu/11274685/Stability_of_Belief_Preface_and_Table_of_Contents_Draft, last access 28th Sep 2016.
- Lepage, F. & Morgan, C. (2003). Probabilistic Canonical Models for Partial Logics. *Notre Dame Journal of Formal Logic*, 44(3), 125–138.
- Lyon, A. (2009). "Philosophy of Probability", in: F. Allhoff (Ed.), *Philosophies of the Sciences: A Guide*, Wiley-Blackwell, UK, 92–125.
- Maher, P. (2006). "The Concept of Inductive Probability", *Erkenntnis*, 65, 185–206.
- Maher, P. (2010). "Explication of Inductive Probability", *Journal of Philosophical Logic*, 39(6), 593–616.

- Novikov, P. S. (2001). "Axiomatic Method", in: M. Hazewinkel (Ed.), *Encyclopedia of Mathematics*: Kluwer Academic Publishers.
- Pearl, J. (1991). "Probabilistic Semantics for Nonmonotonic Reasoning", in: R. C. J. Pollock & R. Cummins (Eds.), *Philosophy and Ai: Essays at the Interface*. MIT press, Cambridge, MA, 157–188.
- Sarkar, S. & Pfeifer, J. (2006). *The Philosophy of Science, an Encyclopedia*. Routledge, Taylor & Francis Group, New York.
- Schilpp, P. A. (1963). *The Philosophy of Rudolf Carnap* (Vol. 11). Open Court, La Salle, Illinois.
- Stalnaker, R. C. (1970). "Probability and Conditionals", *Philosophy of Science*, 37(1), 64–80.
- Van Fraassen, B. A. S. (1995). "Fine-Grained Opinion, Probability, and the Logic of Full Belief", *Journal of Philosophical Logic*, 24(4), 349–377.

Self-Concernment without Self-Reference

Roberto Horácio de sá Pereira

Philosophy Dept.
University of Rio de Janeiro
Av. Pedro Calmon, 550
Cidade Universitária
Rio de Janeiro - RJ, 21941-901
Brasília
robertohsp@yahoo.com.br

Abstract

This paper is a new defense of the old orthodox view that self-consciousness requires self-concepts. My defense relies on two crucial constraints. The first is what I call Bermúdez's Constraint (2007), that is, the view that any attribution of content must account for the intentional behavior of the subject that reflects her own way of understanding the world. The second is the well-known Generality Constraint of Evans (1982), which is also termed the recombability constraint. The claim I want to support in this paper is the following: Since whether and to what extent we can attribute to non-linguistic creatures and prelinguistic infants genuine knowing self-reference or *de se* contents is an open empirical question, the proponents of the nonconceptual self-consciousness face a dilemma. If we are convinced that the available empirical evidence is overwhelming, I argue based on Evans's Generality Constraint—that these self-representations are nothing but primitive prelinguistic self-concepts. However, if we are convinced that the available empirical evidence is not persuasive, I maintain—relying on Bermúdez's Constraint—that we do better by assuming that the subject is not self-represented. The content of her experiences and thoughts are best modeled as simple selfless propositional functions that are true or false relative to the subject of these experiences and thoughts. I refer to this as *self-concernment without self-reference*. Thus, against the recent ingenious work of Peacocke (2014), I claim that there is no compelling reason for postulating nonconceptual middle level self-representation, between self-concernment and conceptual self-reference. However, as I hope to make clear, my claim is quite different from those of other recent oppositions to the idea of nonconceptual self-consciousness. According to the thesis of *self-concernment without self-reference*, the contents of experiences and thoughts are selfless propositional functions, true or false relative to the bearer of the respective mental states.

Introduction

The idea that self-consciousness depends on self-concepts was, until recently, orthodoxy. The best example is found in Baker's paper on this topic (1998). She argues that all sentient beings are *subjects* of experience in the sense that they all experience the world from their own egocentric perspectives. In doing so, they show themselves to be in possession of what Baker calls weak first-person phenomena (Baker, 1998: 60). However, merely being the subject of experiences is not the same as being conscious of oneself *as the subject* of those experiences. Self-consciousness, or what Baker calls strong first-person phenomena, requires the further ability to think of oneself *as oneself*, that is, to conceptualize oneself *as a subject possessing a first-person perspective*. This ability is both a necessary and a sufficient condition for self-consciousness

(i). In this view, self-consciousness is something that only emerges in the course of a long developmental process and that crucially depends upon the acquisition of a self-concept.

This orthodoxy has been challenged from both philosophical (Bermúdez, 1998; Gallagher, 2000; Zahavi, 2006; Peacocke, 2012, 2014) and psychological standpoints (Gallup, 197; Rochat, & Hespos, 1997). The philosophical argument claims that without the postulation of a non-conceptual form of self-consciousness, we cannot avoid an infinite regress (Zahavi, 2006) or defuse a paradox (circularity) in the account of the subject's acquisition of the self-concept that reflects her mastering of the rule of the first-person pronoun (Bermúdez, 1998).

The psychological argument claims that without the existence of primitive forms of self-awareness, one cannot understand the phylogenesis and ontogenesis of the full-fledged linguistic form of self-consciousness. We are told that the empirical findings of developmental psychology, the phenomenological analyses of embodiment, and the studies of pathological self-experience point unequivocally to the existence of primitive forms of self-consciousness that do not require the ability to conceive of oneself as oneself by means of a self-concept. It has become a widespread conviction today that long before the acquisition of a self-concept, conscious beings are already aware of their own selves. According to Gallup (1970), a being is self-aware if it shows the capacity to become the object of its own attention.

The idea of a primitive form of self-consciousness is quite tricky. It might engage us in both a mere verbal dispute around the concepts of self-consciousness and of self-concept and empirical questions as to whether an entity is self-conscious or self-aware. To be sure, these are concepts may be understood quite differently (at least when we compare with what the analytical and the continental traditions have to say about them). Therefore, one might suppose that there is nothing relevant at stake in this new trend in philosophy and psychology (Rochat, & Hespos, 1997). For example, one might suppose that what philosophers and psychologists are calling nonconceptual forms of self-consciousness is simply what Baker calls "weak first-person perspective" ("what is likely to be"). Indeed, to avoid any further ambiguities, several psychologists and philosophers avoid the pompous term "self-consciousness" in favor of the term "self-awareness."

Nevertheless, I claim that beyond any verbal dispute, there is a real philosophical issue at stake here. Regardless of the words used, the question is whether there are conceptual reasons to postulate knowing self-reference (in opposition to an accidental self-reference, when the subjects happens to self-refer without knowing) in case the subject does not possess the self-concept required to specify correctly the alleged *de se* contents of her experiences and thoughts.

Self-Concepts

There are several philosophical routes leading to the idea of self-consciousness without a self-concept. In this section, I first want to consider what Bermúdez calls the deflationary view of self-consciousness (Bermúdez, 1998: 13). This is the orthodox view, according to which the ability to have first-person thoughts is reduced to the ability to employ the first-person pronoun in a way that reflects the subject's mastery of its semantics (ibid.: 15). In this view, a self-concept is just a self-representation that, according to Bermúdez's Constraint, reflects this mastery. Thus, to possess a self-concept is to understand the key rule according to which the subject knowingly refers to herself as the producer of the relevant token "I" whenever she employs that pronoun (Bermúdez, 1998).

With the deflationary view in mind, it is relatively trivial to define, in contrast, nonconceptual first-person, or *de se*, contents: They are any mental state with *de se* contents, even though the subject's self-consciousness does not reflect a mastery the token-reflexive rule of the

first-person pronoun, such as *the producer of that token*. In other words, it does not reflect her knowledge that whenever one employs a token of that pronoun, one refers to oneself as its producer. Let us call this *prelinguistic self-consciousness*.

The first person to raise doubt about the misidentification of nonconceptual self-consciousness with prelinguistic self-consciousness was certainly Meeks (2006). The source of the problem is Bermúdez's so-called *priority thesis*: *Creatures without language are creatures without conceptual abilities* (Bermúdez, 2007: 87). In other words, conceptual abilities are nothing more than linguistic abilities. However, Meeks does not draw the obvious conclusion from Bermúdez's priority thesis: It makes nonconceptualism quite uninteresting. For one thing, the priority thesis trivializes nonconceptualism by rendering almost everything that a nonlinguistic creature represents as nonconceptual content. There is overwhelming empirical evidence in the psychological and ethological literature that supports the assumption that nonlinguistic creatures possess concepts (Herrnstein et al. 1976, Seyfarth et al. 1980; Schütt, 1990; Allen, Bekoff, 1997; Savage-Rumbaugh, & Brakke, 1996). Bermúdez's priority thesis is simply empirically false.

A fortiori, the priority thesis also trivializes the notion of non-conceptual *de se* contents as a prelinguistic self-consciousness. Now any non-linguistic creature or prelinguistic infant is supposedly capable of self-consciousness because they certainly self-refer, although in a way that does not reflect the mastery of the pronoun "I." For example, if I am a creature representing a pond to *my* left or a predator coming toward *me* (Peacocke's examples in 2014: 22), I am in a mental state endowed with a nonconceptual *de se* content. For one thing, under the assumption that I am knowingly self-referring in these cases, my self-representation does not reflect the mastery of the first-person pronoun. How can we *a priori* rule out the existence of non-linguistic self-concepts? Why must we assume that non-linguistic creatures are devoid of self-concepts?

I seriously suspect that the idea of prelinguistic self-consciousness is behind the new trend in developmental psychology that construes pieces of the intentional behavior of animals and prelinguistic infants as nonconceptual forms of self-consciousness on several different levels before the full-fledged linguistic form of self-consciousness. Non-linguistic animals and prelinguistic infants are considered as nonconceptual self-conscious creatures because they represent themselves; however, such self-representations are non-linguistic and do not reflect the mastery of the pronoun of the first person.

The identification of self-concepts with the mastery of the token-reflexive rule of the first-person pronoun raises a further problem. I can only knowingly refer to myself by means of a self-concept that reflects my mastery of the semantics of the first-person pronoun insofar as I know that I satisfy the identifying condition of the token-reflexive rule of the first-person, that is, roughly, as *being the producer of the relevant token*. Again, this trivializes the nonconceptualism content. This renders every nondescriptive reference nonconceptual and every direct, or *de se* non-descriptive self-reference, nonconceptual.

Even worse, the identification of concepts with the mental analogue of linguistic descriptions launches a vicious regress or a vicious circle. As Bermúdez puts it, to self-refer conceptually, one has to master the token-reflexive rule of the first-person pronoun that imposes the identifying condition of *being the producer of the relevant token*. However, the satisfaction of this identifying condition requires another *de se* content to get off the ground. I can only learn that I refer to myself when I employ a relevant token of the first person pronoun, if I *already know* that I am the producer of that token. Bermúdez calls this the "paradox of self-

consciousness" (Bermúdez, 1998: 24), and his way of solving it is to postulate a nonconceptual form of self-consciousness that is prior to and independent of the rule

In fact, there is nothing new in Bermúdez's paradox. The idea has a long philosophical tradition that traces back to Locke and the so-called *Theory of Reflection*. Fichte (1937) was the first to see the problem. If we want to account for self-consciousness as the result of a reflection, we must assume that the subject that carries on the reflection *already knows* that she is the one performing the very act of reflection. In this way, the account presupposes rather than explains self-consciousness. This is what Fichte calls circularity. The alternative is to assume that when the subject performs a higher-order reflection by means of it, she identifies herself as the author of the first-order act of reflection. However, the same question is raised repeatedly, and so we face an infinite regress. Thus, we find ourselves grappling either with a vicious circle or with an infinite regress.

Even though there is deep disagreement about the nature of self-concepts (see, for example, Bermúdez, 1998; Recanati, 2007 and Peacocke, 2014), I think that all parties to the debate generally agree on one basic feature. A self-concept, linguistic or prelinguistic, is the representation of oneself *as oneself* or, in other words, the *knowing* representation of oneself. In that sense, a self-concept is distinct from any other concept that accidentally and unknowingly refers to oneself. Thus, if one sees one's own image reflected in a mirror, but does not recognize oneself, one unknowingly self-refers.

Nevertheless, as a concept, it must fulfill a further constraint. Even though there is deep disagreement about the nature of concepts, I think that all parties to the debate generally agree that states with conceptual content must meet Evans's famous Generality Constraint (1982). In a nutshell, an individual can be credited with the predicative concept *F* should he be able to entertain thoughts in which *F* is applied to any object for which he has individual concepts, such as *b*, *c*, *d* (i. e., *a* is *F*, *b* is *F*, *c* is *F*, and *d* is *F*). Similarly, an individual can be credited with the individual concept *a* should he be able to entertain any thoughts in which *a* is freely recombined with any predicative concept, such as *F*, *G*, and *H* is in her possession (i. e., *a* is *F*, *a* is *G*, *a* is *H*) (Evans, 1982: 104).

Thus, an individual can only be credited with a self-concept if she is capable of freely recombining some knowing self-representation with any predicative concepts in her possession. For example, the self-representation I employ in the thought "I am in pain" is a self-concept because I am able to recombine freely that representation with any such predicative concepts as "I have to see a doctor," "I have to take painkillers," "I must stay in bed," etc. In contrast, a self-representation is supposedly nonconceptual when the subject is unable to recombine it freely with any other predicative concepts in her possession.

That said, all we need to defuse Bermúdez's paradox is to assume, first, that there are primitive non-linguistic self-concepts: knowingly self-representations that meet Evans's generality constraint. Next, that these self-concepts refer directly, that is, without the mediation of any identifying conditions, without the subject's knowledge that she meets any of them. Thus, a prelinguistic infant learns the key token-reflexive rule insofar as she prelinguistically is able to represent herself as the producer of the relevant token. Are there really knowing self-representations that meet Evans's Generality Constraint? This is an empirical question, but data from ethology seem to support it.

Nonconceptual self-consciousness or the weak first-person perspective?

Another route that leads to the idea of self-reference without self-concept is Baker's conception of self-consciousness (1998). According to her, this first ability requires the further higher-order ability of self-attributing first-person thoughts. Jones is not self-conscious when she entertains simple first-order I-thoughts like "I am tall." On Baker's account, Jones only becomes self-conscious when she is able to self-ascribe a first-order, first-person thought, that is, roughly, "I think that I am entertaining the I-thought that 'I am tall'" (Baker, 1998: 330). According to her, a self-concept requires the ability to think or represent oneself as oneself or to conceptualize the distinction between the third and the first-person perspectives. And that (strong first-person perspective) requires in turn the ability to self-ascribe the weak first-person perspective: "I wish I was tall."

A similar idea is suggested by Rosenthal's higher-order theories of consciousness *HOT* (2004). According to him, HOTs (higher-order thoughts about under-order states or thoughts) are already first-person thoughts. The HOT theory postulates that by representing under-order thoughts to which HOTs refer, the under-order thoughts and experiences result in general consciousness. However, HOTs are not self-conscious so long as such they do not represent or refer to the very individual entertaining the HOTs. Thus, self-consciousness emerges only when the HOT disposes the subject of that thought to entertain a further higher-order thought that now identifies her as the individual of the first thought (Rosenthal, 2004: 164).

Even though neither Baker nor Rosenthal are friends of the idea of non-conceptual *de se* contents, with too demanding a conception of self-consciousness in mind, it is relatively trivial to define this notion by contrasting non-conceptual first-person or *de se* contents, while keeping too demanding a conception of self-consciousness in mind. They are any first-person thoughts (e. g., "I am in pain") of a creature that does not possess the further higher-order ability to self-ascribe her own first-person reference ("I think that I am the subject entertaining the first-order thought that 'I am in pain'"). Nor is it about the further ability that disposes her to entertain any higher-order thought that identifies the herself as the individual of the first-order I-thought (e. g., "I think that I am the one who is in pain," etc.).

Nevertheless, as before, these higher-order views on self-consciousness trivialize the notion of non-conceptual *de se* content by rendering any first-order first-person personal thought non-conceptual. To be sure, self-consciousness involves self-reflection (the phenomenological notwithstanding). Still, I see no reason to support the claim that the ability to think of oneself as oneself entails this higher-order ability of self-attributing first-order first-person thoughts. For the same reason I see no reason why should we exclude *a priori* the existence of non-linguistic self-concepts by means of which the subject knowingly represents herself as herself. What is crucial for a self-concept is the satisfaction of Evans's requirement for concept possession (namely, free recombination), rather than higher-order abilities of self-attribution for a first-person perspective.

Let me give you a clear counter-example. Let us suppose a creature is several different phenomenal states just as to be in pain, to be hungry, to be sleepy etc. Then there is something it is like to feel pain, to be hungry etc. for *the creature itself and for nobody else*. Under pressure of social interaction with the caregivers, the creature learns to represent mentally herself *as the one for whom* there is something like it is to be in pain, to be hungry, to be sleepy etc. Thus, when the infant thinks the first-order thought, "I am in pain," she is probably able to recombine freely that self-representation with any other predicative concepts in her possession,

such as “I am hungry,” “I feel sleepy,” etc. (see Evans’s Generality Constraint, 1982: 104). If that is so, then the infant, by fulfilling Evans’s constraint of content attribution, already has a self-concept. Therefore, I see no reason to assume that the infant must possess the further higher-order linguistic ability to self-attribute her own first-person perceptive (“I think that I am entertaining the first-order thought ‘I am in pain’”). Nor must she require the ability to master the token-reflexive rule of the first-person pronoun: I am the producer of a relevant token of the first-pronoun.

Nonconceptual middle level self-representations: The Empirical Dilemma

The intriguing question is why should we accept the existence of middle level non-conceptual self-representations? Peacocke (2014) shows us two independent reasons. Let us consider Peacocke’s own example: “*That a thing (say a predator) is coming towards me; I had such-and-such an encounter; I am moving my head*” (Peacocke, 2014: 6). Peacocke’s first argument is that “it is in the very nature of the type of content in each of these examples that their correctness conditions concerns the subject who enjoys the event” (Peacocke, 2014: 7). The idea is that if we assume that both events have representational contents and further that these contents are best modeled as classical complete propositions by Fregean standards, we can only say that these contents are correct or not under the key assumption that the subject of the events is also represented as an essential part of the content.

The crucial assumption is that we do have to assume that these contents are best modeled as complete propositions. Why, however, need we assume this? Why can we not assume that what is represented by these mental states and events are merely propositional functions, correct or not, of their subject, as in “the pond is to the left of x” or “the predator is coming towards x”? I see no compelling conceptual reasons against the idea of relativist content (see Recanati, 2007; Brogaard, 2012). I return to this question at the end of the paper.

Still, psychologists usually talk here about “self-awareness” as a kind of primitive nonconceptual self-representation (see Gallup, 1970; Rochat, & Hespos, 1997). Nevertheless, the assumption of such “self-awareness” yields more questions than answers. First, is it really compelling to ascribe such self-representations to account for the creature’s behavior? The fact that some creatures recognize their own images in a mirror is good empirical evidence (Gallup, 1970). Still, even if we assume the existence of such self-awareness, a further question arises: Is it a genuine form of self-consciousness when the subject is given in the first-person way? Peacocke answer to this second question is no:

Why is there no such thing as being given as oneself in perception, sensation, or certain kinds of sensory imagination or memory? To ask the question in compressed form and only slightly oversimplified form: why can’t the subject be an object of perception?

However, Peacocke has a second ingenious reason in favor of the middle level non-conceptual self-representations. He argues as follows:

Though a creature can be in a subject-reflexive state that represents it as *F*, and also in a subject-reflexive state that represents it as *G*, nothing in what I have said so far has explained how the subject is in a position to register that he is in a subject-reflexive state of being both *F* and *G*. (Peacocke, 2014: 13f.)

Now subjects must be in fact capable of integrating the contents of those of their conscious states that exhibit subject-reflexivity into such a conjunctive

representation. For a person who possesses and exercises the first-person concept, it is unproblematic how could this be done. A perceptual experience which represents the subject having a pond to his left entitles the subject, other things being equal, to judge the conceptual content of that form *that pond is to left me*. This content contains the first-person, with “me” as the accusative form of the English expression of the first-person concept. The judgment that the pond is left to me is then suitable for inferential integration with other first-person contents, such as I am running strictly straight ahead. (Peacocke, 2014: 14)

In contrast, if both self-representations are non-conceptual, their integration is not supposed to be the product of any conscious inference, but rather results from the fact that they both belong to the same file. While the inferential integration operates at the personal level, the subject’s file on itself does not.

By contrast, as time passes in the first person case, nothing of quite the same kind is required as is needed for the second task in the perceptual case. If a subject at t_1 has a nonconceptual representation of itself a f , by means of a file on itself, it suffices to update this to t_2 to a representation that at the earlier time, it was f . (I continue to use lower case italics for nonconceptual contents.) A past tense predicate capturing this can be combined with other present tense predicates in the subject’s file on itself to yield representations to the effect that the subject as f and is g . (Peacocke, 2014: 16)

Peacocke’s idea of conscious inferential integration satisfies Evans’s Generality Constraint. I conceptually represent myself as being $F \ \& \ G$ as the result of the conscious inference that I conceptually represent myself as being F and that I conceptually represent myself as being G . However, this conscious inferential integration undoubtedly supposes that I am able to recombine freely my own self-representation with any other predicative concepts in my possession in the first place: H, J, R , etc. or I am H ; I am J ; I am R , etc.

However ingenious, Peacocke’s explanation is far from convincing. For one thing, his own examples are centered on the simple cases of the inference of conjunction-introduction of properties that the subject self-ascribes, and the only explanation of non-inferential integration he provides is the update of the same property. Let us reconsider Peacocke’s other example of middle level self-representation: “That predator is coming towards *me*” (Peacocke, 2014: 6). To begin with, if the creature possess mental states that refer to itself, such as *that predator is coming towards me*, it is reasonable to assume to she also possesses other subject-reflexive states that represent itself, such as *I am in danger*, *I must space myself from the predator*, etc.

Therefore, the same empirical evidence that might convince us to attribute to a creature *de se* contents, such as *that predator is coming towards me*, also allow to attribute to the same creature several similar I-thoughts, in which self-representation is freely re-combined with any other predicative concepts that she possesses: “I am in danger”; “I must try to escape; run”; and so on. Thus, Peacocke’s own examples of middle level self-representations easily meet Evans’s General Constraint for concept attribution. They seem to be primitive self-concepts rather than nonconceptual self-representations. Peacocke says:

(..) States of consciousness with nonconceptual *de se* contents such as *there is a pond to my left* and *I am running straight ahead* cannot be reached on the model of conceptualized inferences of conjunction-introduction. (Peacocke, 2014: 29)

Furthermore, if we consider the creature's intentional behavior as good empirical evidence from which to attribute to the creature *de se* contents, such as *there is a pond to my left* and *that I am running straight ahead*, we cannot help but fill the gap and attribute to it the further *de se* content *I am thirsty*. The natural assumption here is to regard the same overall empirical evidence as good reasons to consider the creature's intentional behavior *of running straight ahead to the pond* as the result of an inference. Roughly, *the pond is to the left of me; I am thirsty; therefore, I am running quickly, straight ahead to pond to drink water*. The natural assumption here is that representations can only be integrated non-inferentially and sub-personally, if the subject is not part of the putative *de se* content, but rather merely self-concerned by the selfless content of her visual experience.

Perry's notion of unarticulated constituents

In his attack on the idea of nonconceptual self-consciousness, Meeks (2006) was the first scholar to appeal to Perry's famous thought-experiment (1986/2000). Considering it, Meeks claims, "we cannot extend the immunity condition to account for the ostensibly self-conscious states we may wish to ascribe to such creatures (that lack self-concept)" (Meeks, 2006: 97). Musholt (2013), inspired by Meeks, argues that "the nonconceptual representational contents of perception and bodily experience are neither self-representational, nor do they fall under the category of representations that can be said to be immune to error through misidentification" (Musholt, 2013: 23). However, before proceeding, it is worth mentioning, first, how Perry himself describes his thought-experiment

Perry (1986/2000) invites us to consider Z-landers, a group or a tribe that lives in complete isolation and that has never left Z-land, its present place of residence. What matters to us is the following. When residents of Z-land file weather reports like "it is raining," "Z-land" has an argument role of a certain relation $\langle \text{rains}; \text{Z-land} \rangle$. The correct conditions of its content certainly *involve* Z-land, the place where the Z-landers' weather report is filed. That content is correct or accurate if it is raining in Z-land at the time Z-landers report this weather condition. However, as Z-land has argument role that never changes, Z-landers do not need to worry about Z-land. According to Perry, Z-land is a so-called "unarticulated constituent" of the weather report "it is raining"; that is, it is a constituent of their report that is neither verbally articulated nor mentally represented by their utterances.

Let us suppose now that anthropologists find Z-land. As usual, an exchange of gifts takes place, and residents of Z-land receive cell phones from the anthropologists to communicate with their new fellows outside Z-land. Now things change. When they communicate weather conditions in Z-land to the anthropologists outside of Z-land, they must learn to articulate Z-Land in their weather reports. They thus acquire the key concept "Z-land."

The analogy to the problem of self-reference is straightforward, to the extent that the non-linguistic animal and the prelinguistic infant are just an egocentric, unchanging frame of reference in the subject's experiences. They are also an argument role that never change; therefore, in these states, the subject does not have to worry about herself when she experiences or thinks something. Perry's assumption is that the subject, as the egocentric frame of reference, is also an unarticulated constituent of the content of her visual experience. Things naturally change when the prelinguistic infant begins to acquire language and starts to communicate her experiences and thoughts to her parents or to her caregivers. Now the subject of experience becomes an argument role that changes constantly. Thus, the infant must learn to articulate her self-concept in her mental states reports.

However, in his seminal paper of 1986/2000, Perry is ambiguous about his own notion of “unarticulated constituent”. In a few passages, he clearly states that neither Z-land nor the subject are *referred to* as part of the content itself:

Let us develop a little more vocabulary to mark this distinction. We shall reserve “about” for the relation between a statement and the constituents of its content, articulated and unarticulated. We shall say a belief or assertion *concerns* the objects that its truth is relative to. So the Z-lander’s assertions and beliefs *concern* Z-land, but are not *about* Z-land. (Perry, 1986/2000: 179)

In other passages like this one, Perry *seems* to say that both Z-land and the self, as “unarticulated constituents”, are *not referred to* as part of the content of their respective experiences and thoughts, but only *concerned* with the same experiences and thoughts (actually, that is the view that Recanati has been defending for almost a decade, since 2007). However, in his same seminal paper of 1986/2000, he clearly seems to state the opposite:

The unarticulated constituent is not designated **by any part** of the statement, but it is identified by the statement **as a whole**. The statement is *about* the unarticulated constituent, as well as the articulated ones. So, the theory is (i) some sentences are such that statements made with them are about unarticulated constituents; (ii) among those that are, the meaning of some requires statements made with them to be about a fixed constituent, no matter what the context; whereas (iii) others are about a constituent with a certain relationship to the speaker, the context of use determining which object has that relationship. (Perry, 1986/2000: 174; emphasis in bold is mine)

Despite all appearance to contrary, Perry’s official doctrine after 1986/2000 is that the “unarticulated constituent” of the content is certainly referred to by the subject’s entire mental state, even though it is not mentally or verbally articulated in utterances. The reason is clearly articulated as follows:

Similarly, the Z-landers’ beliefs about the weather lead them to actions that make sense if it is raining in Z-land. So, it seems that those beliefs ought to be true, depending on how the weather is in Z-land. *And so it seems that the objects of the belief should be about Z-land, so that they will be true or false depending on the weather there.* (Perry, 1986/2000: 214; emphasis is mine)

Thus, without the key concept “Z-land,” the Z-lander’s weather reports *as a whole* undoubtedly refer to Z-land as an unarticulated constituent of their content. Likewise, without a self-concept the subject’s experiences and thoughts undoubtedly refer to the subject of those experiences as an unarticulated constituent of the *de se* content of her experiences and thoughts. Perry supports this claim by arguing that otherwise those contents would be an incomplete, in the sense of being a propositional function without a determined truth-value.

If this is Perry’s conception, Meeks misunderstands his position when he says, “the Z-landers’ weather reports [. . . neither explicitly nor implicitly represent Z-land and are therefore not about it” (Meeks, 2006: 95) and adds:

In the case of proprioception, then, such states represent the properties and states of one’s body without representing oneself, instead simply concerning oneself in that they regulate and mediate one’s own behavior in the appropriate way. We may need to identify the subject of such states when specifying the conditions under which such states successfully represent (or misrepresent)

sent) the property or state in question, but the states themselves need not represent the proprioceiving subject at all. (Meeks, 2006: 95)

To be sure, without a self-concept *no part* of the subject's statements or thoughts refers to the subject. Nonetheless, as Perry clearly puts it (quote above), the unarticulated constituent (the subject) is not designated or referred to by *any part* of the statement, but by the statement or the thought *as a whole* (Perry, 1986/2000: 174)

Regardless of whether Meeks case is based on a misunderstanding of Perry's official doctrine, his explanation of why one cannot extend immunity through misidentification is also not satisfactory. Meeks complains that states that are immune to error through misidentification require complex structured conceptual contents (Meeks, 2006: 98). Because of its noncompositionality, nonconceptual content cannot accurately represent the subject of a self-ascription while misrepresenting the property; it can only misrepresent tout court (or else fail to count as genuine content). To be sure, nonconceptual content is noncompositional, otherwise it would satisfy Evans's Generality Constraint (Evans, 1982: 104). Still, it does not seem to follow that nonconceptual states *can only misrepresent tout court*. On the contrary, it seems to me quite possible to misrepresent, say, the color of this object, while being immune to error through the misidentification of *this object*.

In contrast, Musholt argues that the notion of immunity to error through misidentification cannot apply to nonconceptual content in the first place. For one thing, the immunity to error through misidentification can only arise at the level of judgment, not at the level of nonconceptual content (Musholt, 2013: 19). According to her, "it is a category mistake, so to speak, to try to apply the notion of immunity at the level of nonconceptual content" (Musholt, 2013: 19). To be sure, judgments are the paradigmatic cases of immunity to error through misidentification. Still, I see in this no reason contrary to the assumption that when I nonconceptually represent *that color* in normal conditions, I am also immune to error through the misidentification of the object that I mentally demonstrate as "that" while misrepresenting its color.

The problem is not that of extending the notion of immunity to error through misidentification to nonconceptual contents in general. Rather, it is that of extending that notion to the idea of nonconceptual self-consciousness in particular. For one thing, immunity through error misidentification is a limiting case, where the reference dispenses identification of the referent. However, if the subject of exteroception and proprioception never self-refers, it is difficult to understand how proprioception could be immune to misidentification in the first place. Still, we must further assume that even without self-reference, any experience with a phenomenal character provide self-specifying information whose source is the subject: there is something that is like to be in a phenomenal state *for the organism*. Thus, when the subject begins to self-refer knowingly, the self-reference is immune to misidentification because it is based on this intrinsic relation between the phenomenal states and the concerned subject of those states.

Nevertheless, let us go on. Musholt's appealing to Perry's thought-experiment suffers from the same ambiguity of Perry's seminal paper. Sometimes, she seems to merge both readings of the notion of unarticulated content in just one: "The squirrel representation does not need to be *about* itself, it does not need to contain a *self-referring component* in order to be action-guiding" (Musholt, 2013: 10–11). Now, from the fact that mental states do not contain particular components to refer to the subject, it does not follow that Perry does not regard the subject as implicit self-referred.

Elsewhere, Musholt seems to oscillate between the two readings. Like Meeks, she indicates in several passages that Z-land is not a matter of reference but merely one of concern:

Z-landers' thoughts about weather *concern* Z-land insofar as they lead to behavior that is appropriate to the weather in Z-land (e. g., taking an umbrella when leaving the house upon thinking "It is raining"), but Z-land does not have to be represented for this to hold (hence their thoughts are not *about* Z-land). (Musholt, 2013: 12)

Nevertheless, in other passages she clearly assumes that Z-land is part of the content of the Z-landers' weather reports, as when, for example, she states:

Z-land figures as an "unarticulated constituent" of the utterance because in order to determine the truth conditions of the sentence "It is raining" we need a location (in this case Z-land)—the sentence will be true if it is indeed raining in Z-land. (Musholt, 2013: 11–12)

However, when, in a footnote, Musholt clarifies her opposition to explicit self-representations and implicit self-related information, she leaves no doubt that she is assuming Perry's official doctrine of the unarticulated constituents:

A fact or state of affairs is represented *explicitly when the mental state in question contains a component that directly refers to this fact or states of affairs*. In contrast, a fact or state of affairs is *implicit* in a mental representation when the mental state in question does not contain a component that directly refers to this fact, but *when this fact or state of affairs is conveyed as part of the contextual function of the mental state*. (Musholt, 2013: 9; my emphases)

Nevertheless, if the same fact that is explicitly represented by a mental state and implicitly conveyed by the context, Musholt's entire case against nonconceptual self-consciousness collapses. She argues, "theories of nonconceptual self-consciousness are incomplete insofar as they only establish the existence of *implicit* self-related information in perception and proprioception, but not the existence of *explicit* self-representation" (Musholt, 2013: 8). However, the question is why the proponents of the nonconceptual self-consciousness need to assume that there is an explicit nonconceptual component in the mental states of non-linguistic creatures and prelinguistic infant that refer to themselves. All they need is the acknowledgement of implicit, self-related information that indicates the presence of the subject in the *de se* content of her own exteroceptive and proprioceptive experiences, without self-concepts or "explicit self-representation".

The only way to build a case against the idea of nonconceptual self-consciousness, based on the Perry's official notion of unarticulated constituent, is the following. Before the acquisition of the key concept "Z-land," Z-lander's weather reports as a whole already designate Z-land as a part of their content, otherwise the content would be *a mere propositional function without a fixed truth-value*. Likewise, without the acquisition of the key self-concept, the prelinguistic infant's thoughts and experiences as a whole already designate herself as a part of the content, otherwise the content would be a mere propositional function without a fixed truth-value. Now, the opponents of the idea of nonconceptual self-conscious could argue as follows: To be sure, even without a self-concept the prelinguistic infant and other nonlinguistic creature's thoughts and experiences already designate the infant herself. However, without the key self-concept, her self-reference is *unknowing*. Now since self-consciousness is knowing rather than accidental self-reference, the prelinguistic infant may be self-represented by her experiences and thoughts, but she is not genuinely self-conscious.

Now, this line of thought clearly supposes what in the literature (Heck, 2000) is known as the state nonconceptualism (state view). According to the state view, nonconceptualism is a property of mental states, that is, a view about the relation between the subject undergoing a mental state and the representational content of that state. A mental state is state-nonconceptual when it is concept-independent. Conversely, a mental state is state-conceptual when the subject cannot be in that the state in question without possessing the concepts involved in the correct specification of its contents. Thus, according to the state view, the main difference between nonconceptual and conceptual states is that only in the second case does the subject knowingly refer, that is, understand to what his mental state refers. Therefore, according to state nonconceptualism, experiences and attitudes might share the same content, even when the subject is in different types of states.

In contrast, according to the nonconceptual content view, nonconceptualism is better characterized in terms of *the kind of content* that experiences possess, as opposed to the content of beliefs and other propositional attitudes. A mental state is content-nonconceptual when the content of the state is of a particular type, namely, when it is not composed of concepts. Conversely, a mental state is content-conceptual when it is a structured complex compounded of concepts. Therefore, according to content nonconceptualism, experiences and propositional attitudes could not possibly share the same representational content.

Accordingly, before and after the acquisition of key concepts, Z-landers and prelinguistic infants' mental states *as wholes* do designate Z-land and the infants themselves, respectively. The contents of their states may be modeled Russellian propositions consisting of the very designated entities, such as <Z-land, the property of being raining> and <subject, the property of being in pain>. According to the state view, the only difference is that in both cases, without the key concepts, they have only the faintest idea of what the whole mental states represent.

However, here a crucial asymmetry emerges between Z-landers and prelinguistic infants. By assuming that Z-landers already refer to Z-land by the weather reports as a whole, without the relevant concept "Z-land," Z-landers do represent Z-land, albeit nonconceptually. In contrast, as we saw, by assuming that the prelinguistic infant can already self-refer as an unarticulated constituent of the content of her experiences without the relevant self-concept, we cannot talk about nonconceptual self-consciousness because self-consciousness is knowingly self-reference, that is, a non-accidental self-reference when the subjects knows that she self-refers. In other words, without a self-concept, her self-reference is what psychologists like Gallup (1970) called *self-awareness*: an accidental self-reference. The assumption here is that the creature sees itself in the mirror, but without a self-concept she does not actually recognizes herself as Gallup claimed.

Self-Concernment without Self-Reference

Now, based on Bermúdez's constraint of content attribution (2007), I want to present and defend my own view, which I call *self-concernment without self-reference*. Against Perry and all his followers, I will argue that without the key concept "Z-land" and the self-concepts, what is missing is not a knowing reference (Z-land) or a knowing self-reference, that is, a non-accidental self-reference (prelinguistic infant). What is missing is reference and self-reference in the first place! However, to avoid the ambiguity I found in Meeks (2006) and Musholt (2013), I claim additionally that without self-concepts, the contents of creatures' states are propositional functions that are true or false relative to the bearers' of those states.

Thus, the crucial question we have to face is whether the state view can satisfy Bermúdez's constraint of content-attribution. This is a reasonable view, according to which any attribution

of content must be the best available account for the subject's intentional behavior that reflects her way of understanding the world (Bermúdez, 2007). However, what content we should attribute to a creature in the face of her intentional behavior is an open empirical question, which is not up to us, as philosophers, to decide. Consequently, the questions are the following: Can we really say that the content of Z-landers reports remains unchanged before and after they learn the concept "Z-land"? Before learning the concept "Z-land," do Z-Landers possess the ability to refer to Z-land as the state that nonconceptualism supposes?

The only reason in Perry's paper that supports his view is clearly that Z-land must figure as an "unarticulated constituent" of the utterance, since, otherwise, we do not have a complete proposition, with a fixed truth-value, but rather a propositional function that is true or false relative to Z-land.

This last assumption is also questionable, however (Recanati, 2007; Brogaard, 2012, etc.). Within the framework of Kaplanian semantics (1989), a sentence *S* is true at a context of use *c* if the proposition *p*, expressed by *S* at *c*, is true at the default circumstance of evaluation, determined by *c*. The default circumstances of evaluation are pairs of a world and a time, so a proposition *p* is true at a given circumstance if the proposition is true at the world and time of that circumstance. Nevertheless, nothing hinders us to enlarge these circumstances, including the locale and the subject.

In this relativistic framework, the natural assumption is to think that content of Z-lander's weather reports as simple propositional functions that are true or false relative to Z-land (the argument). I see no compelling conceptual reasons against such a suggestion. However, I cannot defend such a relativistic claim here for obvious reasons of space. The best defense that I know is Brogaard's (2012).

Interestingly, even Perry seems to think that a propositional function could do the job of making sense of the Z-lander's reports and actions:

The only job of their assertions and beliefs concerning the weather is to deal with the nature of the weather in Z-land. Their assertions and beliefs are satisfactory, insofar as their "weather constituent"—rain, snow, sleet, etc.—matches the weather in Z-land, were our need also to register the place of the weather. By taking the propositional content of their beliefs to be propositional functions, rather than complete propositions, and taking them to be true or false relative to Z-land, we mark this difference. (Perry, 1986/2000: 215)

Nevertheless, Bermúdez's Constraint is a powerful reason against Perry's idea that "the argument role that never changes" is an unarticulated constituent of the content. In accounting for the Z-lander's communicative exchanges about the weather in Z-land, we do better in assuming that they are not referring to Z-land. For one thing, residents of Z-land, who have never left their country, cannot *discriminate* Z-land from other lands; they cannot *visually indicate*, *track*, or *pick out* Z-land on a map. Therefore, they cannot refer to Z-land not, even by means of the concept "here." If by chance they possess the concept, "here" refers at most to a certain place *inside* Z-land, rather than Z-land itself. For one thing, since they never left Z-land, they have no other land in mind to oppose to "here"

Compare this to a nonconceptual representation of something in my visual field. Being nonconceptual, I really do not know or understand what I am seeing. Still, I can easily discriminate it from other objects in the field. I can easily indicate it by a pointing gesture and the pronoun "this" or by the adverb "here." I can easily track it down as it moves. None of these abilities is available to the residents of Z-land.

Perry's own examples further substantiate the same point. Time zones certainly have argument roles in any time report. However, before the Europeans' great discoveries of new continents, time zones had argument roles in time reports that never change. Thus, in the light of Bermúdez's constraint of content attribution, people never refer to time zones as unarticulated constituents of their time reports, because they do not have the ability to discriminate times zones or to indicate or pick out a particular time zone. Now on Bermúdez's Constraint, the most parsimonious account of the Z-landers' weather report is to assume that Z-land is a mere aspect of the wide circumstance of evaluation rather than an unarticulated constituent of the placeless content itself.

In this regard, the reference to Z-land as the reference to a time zone is quite different from the reference to objects and properties within the subject's perceptual field. For one thing, like entities postulated by science (quark, atom, energy, photon, etc.), Z-land is never given as *an object* of perception that the residents of Z-land can discriminate from other places outside Z-land. Imagine the first man who arrived at the idea of "universe." In these cases, references rely on and are created by concepts. Now, after their first acquaintance with the anthropologists, we must assume that residents of Z-land begin to refer to Z-land, since that assumption is the best available explanation for the Z-landers' intentional behavior of communicating with their new friends (anthropologists) outside Z-land that reflects their way of grasping the world.

As before, the analogy to the problem self-reference is straightforward. To the extent that the subject is just an egocentric frame of reference in her experiences that never changes (she is also "an argument role that never changes"), the infant has no mental abilities to discriminate herself from others that could justify a self-reference as an unarticulated constituent of her experience. In the particular case of self-reference, the ability to discriminate herself from the others come together with the ability to knowingly represent oneself as such.

Thus, the best available account of the content of the prelinguistic infant's mental states is the assumption that it takes the forms of selfless propositional functions that are true or false of the subject of the mental states. This is what I am calling here *self-concernment without self-reference*. The subject is *concerned* by what her mental state represent (a propositional function) insofar as she belongs to the wide circumstance of evaluation of that content. But she is *not referred to* the extent that she is a constituent of the content. Let us suppose our subject sees a predator coming in her direction. If she possesses the concept of a predator, she might think: "That predator is coming (towards)". This propositional function is true or false, relative to the subject of that mental state or event. In that sense the subject is merely concerned by her content rather than being represented.

Conclusion

Even though we are arguing against Perry's notion of unarticulated constituent of the content on the basis of Bermúdez's constraint of content attribution, we must recognize that it is an open empirical question whether we should consider the prey's pieces of intentional behavior as compelling empirical evidence for attribution of a self-representation to non-linguistic and prelinguistic infants.

Therefore, we face the same dilemma as before. If we think that it is not, the most parsimonious explanation for the creature's intentional behavior that reflects her comprehension the world, is to assume that the content of her experience is a mere propositional function that is true or false relative to the subject of the mental state. In order to account for why the prey tries to escape, we do not need to assume that she is representing herself knowingly as herself.

By contrast, if we may consider the prey's intentional behavior as compelling empirical evidence for the attribution of a self-representation, the subject is part of the *de se* content of her visual experience. Nonetheless, as we saw, if that creature's intentional behavior is compelling empirical evidence to attribute to her the first-person "a predator is coming towards *me*," it must also be seen as compelling empirical evidence to attribute to the creature several similar I-thoughts in which the self-representation is freely recombined with any other predicative concepts that she possesses ("I must run," "I must climb the nearest tree," "I am about to die," etc.). Thus, the self-representation easily meets Evans's constraint for concept attribution. By all accounts, they are self-concepts rather than non-conceptual self-representations.

However, as in case of Z-landers, things change when the subject starts to communicate with someone who holds a quite different viewpoint than hers. Let us suppose a fellow addresses her, saying, "My dear, I am without my glasses. Is that a predator coming?" Now the subject cannot help but refer to her own viewpoint in response to this question: "I can assure you that a predator (a lion) is coming." It is no longer an open empirical question as to whether we should consider the prey's behavior as compelling empirical evidence for the attribution of a self-representation to the prey. In the face of the communicative exchange, the subject is no more an aspect of the circumstance of evaluation, but rather an essential part of the *de se* content itself. The best available explanation for her intentional behavior that reflects her way of perceiving the world is to assume that she is essentially part of the content. Still, by all accounts, the subject's self-representation meets Evans's Generality Constraint on concept attribution and hence qualifies as self-concept. Thus, we come to the following skeptical conclusion: There is no compelling reason to accept a middle level nonconceptual self-representation.

References

- Allen, C. & Bekoff, M. (1997). *Species of Mind. The Philosophy and Biology of Cognitive Ethology*. MIT Press, Cambridge.
- Baker, L. R. (1998). 'The First-Person Perspective: A Test for Naturalism.', *American Philosophical Quarterly* 35(4), 327–348.
- Bermúdez, J. L. (1998). *The Paradox of Self-Consciousness*. MIT Press, Cambridge, MA.
- Bermúdez, J. L. (2007). 'What Is at Stake in the Debate on Nonconceptual Content?' *Philosophical Perspectives* 21, 55–72.
- Brogaard, B. (2012). *Transient Truths: An Essay in the Metaphysics of Propositions*. Oxford University Press, Oxford.
- Evans, G. (1982). *The Varieties of Reference*. Oxford University Press, Oxford.
- Fichte, J. G. (1937). *Nachgelassene Schriften. Band 2: Schriften aus den Jahren 1790-1800*, Jacob, H. (ed.), Junker und Dünhaupt,
- Gallagher, S. (2000). 'Philosophical Conceptions of Self: Implications for Cognitive Science.', *Trends in Cognitive Sciences* 4(1), 14–21.
- Gallup, G. Jr. (1970). 'Chimpanzees: Self-Recognition.', *Science* 167, 86–87.
- Heck, R. (2000), 'Nonconceptual Content and the "Space of Reason,"', *The Philosophical Review* 109, 483–523.
- Herrnstein, R. J., et al. (1976). 'Natural Concepts in Pigeons', *Journal of Experimental Psychology, Animal Behavior Processes* 2, 285–302.
- Kaplan, D. (1989). Demonstratives: An Essay on the Semantics, Logic, Metaphysics, and Epistemology of Demonstratives and Other Indexical" i Almog, P. & Wettstein, H. (eds.), 'Themes from K' Oxford University Press, New York, pp. 481–563.

- Meeks, R. (2006). 'Why nonconceptual content cannot be immune to error through misidentification', *European Review of Philosophy* 6, 83–102.
- Musholt, K. (2013). 'Self-consciousness and nonconceptual content.', *Philosophical Studies* 63 3), 649–672.
- Peacocke, C. (2012). 'Subjects and Consciousness' Coliva, A. (ed.), 'Self and Self-Knowledge' Oxford University Press, Oxford, pp. 74–101.
- Peacocke, C. (2014). *The Mirror of the World. Subjects, Consciousness, and Self-consciousness*. Oxford University Press, Oxford.
- Perry, J. (1986). 'Thought Without Representation.', *Aristotelian Society Supplementary Volume*, 137–52. eprinted and quoted from: Perry, J. (2000). *The Problem of the Essential Indexical and Other Essays*, Oxford University Press, Oxford, 1993. *New, Expanded Edition*. Stanford University, CSLI, 2000. Center
- Recanati, F. (2007). *Perspectival Thought: A Plea for (Moderate) Relativism*. Oxford University Press, Oxford.
- Rochat, P. & Hespos, S. J. (1997). 'Differential Rooting Response by Neonates: Evidence for an Early Sense of Self.' *Early Development and Parenting*, (3/4), 105–112.
- Rosenthal, D. M. (2004). 'Being conscious of ourselves', *The Monist* 87(2), 161–184.
- Savage-Rumbaugh, S. & Brakke K. E. (1996). 'Animal Language: Methodological and Interpretative Issues' in Bekoff, & D. Jamieson, D. (eds.), 'Readings in Animal Cognition', MIT Press, A, pp. 269–288.
- Schütt, H. P (1990). *Die Vernunft der Tiere*. Keip Verlag, Frankfurt.
- Seyfarth, R. M., et al. (1980). 'Vervet Monkeys Alarm Calls: Semantic Communication in a Free-ranging Primate', *Animal Behaviour* 28, 1070–1094.
- Zahavi, D. (2006). 'Thinking about (Self-) Consciousness: Phenomenological Perspectives' in Kriegel, & Williford, K. (eds.), 'Self-Representational Approaches to Consciousness', MIT Press, Cambridge 273–295.