A Typology of Conceptual Explications

Dirk Greimann

Fluminense Federal University

A Typology of Conceptual Explications

Dirk Greimann
Fluminense Federal University

Conceptual explication is an art having approximately the same importance for philosophy that logical deduction has for mathematics. To be sure, conceptual explication is important, not only for philosophy, but also for any other science. For all sciences are committed to clarify their own concepts. But philosophy is tied to conceptual explication in a special way, because the philosopher's work consists in large part in explicating concepts, just as the mathematician's work consists in large part in deriving theorems. For this reason, philosophy has closer ties with conceptual explication than does any other science.

With regard to logical deduction, we already have elaborate theories at our disposal that explain the rules we follow implicitly when we are making logical deductions. Unfortunately, this does not apply also to the explication of concepts. On the contrary, it seems that Carnap's diagnosis of the state of the art from 1950 is still valid, according to which "(p)hilosophers, scientists and mathematicians make explications very frequently", but "they do not often discuss the general rules which they follow implicitly" (Carnap 1950, p. 7).

In what follows, my aim is to make some steps toward the construction of a theory of conceptual explication whose tasks are basically the following four: (i) the explication of "explication"; (ii) the distinction of the main types of explication; (iii) the formulation and justification of criteria of adequacy for each type; (iv) the construction of a system of rules for each type whose observation guarantees the adequacy of the corresponding conceptual explications. Such a theory would not only contribute to the methodological self-awareness of philosophy, but it may also help to resolve some of

the confusions that have already crept into the practice of conceptual explication. Thus, Richard Kirkham has persuasively shown that the current debate about the concept of truth suffers from a "multidimensional" confusion that results from mixing up different projects of explication with different goals and different criteria of adequacy.[1] The clear distinction of these projects and the corresponding types of explication is perhaps the most important desideratum of the current theory of truth.

For this reason, I shall focus in this paper on the distinction of the various types of conceptual explications.[2] The main goal is to set up a typology (or "taxonomy") of conceptual explications. The paper is structured as follows. In section 1, the concept of "conceptual explication" is briefly explicated. Section 2 is devoted to the task of explaining the criteria of classification for conceptual explications on which the typology is based. In sections 3 to 7, the main types of conceptual explication are distinguished and the corresponding criteria of adequacy are sketched. Finally, in section 8, the adequacy of the proposed typology is briefly discussed.

1 Explication of "explication"

In the history of philosophy, conceptual explications have been made partly with a descriptive and partly with a revisionary intention.[3] The aim of a descriptive explication is to describe the content and function of a given concept as faithfully as possible. A revisionary explication, on the other hand, aims to transform a concept that is useful but somehow defective from a scientific point of view into a scientifically more respectable concept. Such defects may be, e.g., the vagueness of a given concept, its inconsistency or the lack of a proper principle of individuation for it. The linguistic theories of truth such as the disquotation theory and the minimalist theory are examples of descriptive explications, and the epistemological theo-

---

[1] Cf. Kirkham 1995, chapter 1.

[2] In Geimann (2007), I concentrate on the methodological rules for the correct explication of concepts.

[3] This distinction is analogous to Strawson's well-known distinction between descriptive and revisionary metaphysics.

ries like the consensus and the verification theory are examples of revisionary ones.

Following Quine, it has become common practice in large parts of Analytic Philosophy to call descriptive explications "conceptual analyses" and revisionary explications simply "conceptual explications".[4] This terminology may be legitimate, but it is certainly not appropriate with regard to our ends. Since we want to construct a classification of conceptual explications, we need a more general notion of explication that allows us to consider descriptive and revisionary explications as special cases. To take this into account, I shall follow here the traditional approach to conceive of an explication as an answer to a question of the form, "What is X?" like "What is truth?" and "What is knowledge?" A descriptive explication of truth, for instance, explains what is actually understood by "truth" in natural language, and a revisionary one what should be understood by "truth" in science.

According to Quine, the essence of conceptual explication consists in the elimination of the explicandum, that is, its substitution by a similar but scientifically more respectable concept. He in fact defines: *explicare est eliminare* (1960, § 53). Even when we restrict this explication of "explication" to the revisionary case, it does not appear to be satisfactory. For, in order to transform a defective concept into a scientifically more respectable one, it is neither possible nor necessary in all cases to replace it by another one. Take, for instance, the pretheoretic concept of set. From the point of view of set theory, this concept is vague and ambiguous. But, to overcome these defects, it is not possible to substitute the explicandum by another concept, because the concept of set is primitive. Nor is this necessary, because the defects can be overcome by means of an axiomatic characterization or kindred methods.

Moreover, Quine's explication of "explication" does not seem to be materially adequate, for two reasons. First, the explicandum and the explicans are not sufficiently similar: to "explicate" does not mean to "eliminate", but to make something explicit. Second, this explication does not take into account that the majority of the philosophical concepts are "primitive". It is true that in set theory and

---

[4] Cf. Quine 1960, §§ 39, 53.

other formal disciplines the conceptual explications actually given consist in the elimination of the explicanda, that is, more precisely, in their reduction to a few basic concepts. But the same does not apply to philosophy, because the main philosophical concepts (such as existence, truth, meaning, freedom and morality) cannot be reduced to more fundamental concepts.[5] The elimination of the explicanda is hence a wrong ideal in philosophy.

What we have called here "conceptual explication" is frequently called "definition" in the literature.[6] Unfortunately, the term "definition" is extremely vague and ill-defined. A definition in Frege's sense consists in the logical decomposition of a complex sense into its constituents or in a stipulation to the effect that the definiendum has the same sense as the definiendum.[7] His thesis of the indefinability of truth says that truth is a logically simple (atomic) concept. A definition in Tarski's sense is something quite different. To define X in his sense means to formulate an open sentence whose extension is X, where X may be an individual or a set. His theorem of the indefinability of truth says that in theories whose language is semantically closed it is impossible to construct in a consistent way an open sentence whose extension is the set of all true sentences. To say that truth is indefinable in this sense is to say that truth is inexpressible.[8] Other authors mean by "definition" the reduction of the definiens to more fundamental concepts. When, for instance, Davidson speaks of the "folly of trying to define truth", what he has in mind is the reduction of truth to more fundamental concepts.[9] In order not to complicate things unnecessarily, I shall not try to define "definition" here, but to avoid using this term.

---

[5] This is confirmed by the fact that no philosophical concept that has been successfully reduced to other concepts until now.

[6] See, for instance, Gupta 2008.

[7] Cf. Frege 1914, pp. 226-228.

[8] Cf. Tarski 1933.

[9] Cf. Davidson 1996.

2 Criteria for distinguishing types of conceptual explication

When we want to construe a typology of conceptual explications, we need criteria of classification that enable us to distinguish between different types of conceptual explications. Such criteria may refer either to the purpose of an explication, or to the method used to realize the explication, or to the kind of its explicandum, or to its criteria of adequacy.

In the literature it is common to distinguish between three basic types of explications with regard to their criteria of material adequacy: descriptive, revisionary and stipulative explications. A descriptive explication is materially adequate only if it is faithful to the original meaning of the explicandum, i.e., the explicandum and the explicans must have the same meaning. In the case of revisionary explications, by contrast, the explicans and the explicandum are not supposed to have the same meaning, because otherwise the explicans could not be transformed into a scientifically more respectable concept; you cannot wash the fur without making it wet. Nevertheless, in the revisionary case the explicandum and the explicans are supposed to have similar meanings, because otherwise the explicans could not substitute the explicandum. Finally, in the case of stipulative explications, the explicandum and the explicans need not even have similar meanings, because everyone is free to stipulate what he wants to understand by the explicandum.

A second classification refers to the method used in explications. In this respect, we may distinguish between explicit and implicit definitions, ostensive and lexical definitions, operative and contextual definitions, axiomatic characterizations, Fregean elucidations, ramseyfications etc.

The third criterion of classification refers to the kind of the object of explication, that is, the type of the explicandum. In this respect we may distinguish between extensional and intensional explications, and between nominal and real explications.

Finally, the fourth criterion of classification refers to the purpose of explications. Thus, descriptive and revisionary explications differ, in the first place, with regard to their purposes. The same applies to the distinction between stipulative and descriptive explications.

It is clear that among the four criteria mentioned the last two are the most important ones. For, the methods of explications and their criteria of adequacy depend on their objects and their purposes. It can, moreover, be shown that a larger part of the confusions to be found in the literature about the explication of truth refer to the objects and the purposes of explications. For these reasons, I shall concentrate in what follows on the classification of explications with regard to their objects and their purposes.

3 Real and nominal explications

When we want to answer the question "What is water?", we may consult either a linguist or a chemist. The linguist will explicate the linguistic meaning of "water", that is, the criteria of the correct application of the word "water". Part of his explication may be that "water" applies to those entities that are liquid, colourless, odourless, can be used to satisfy one's thirst, and so on.[10] The explicandum is in this case the descriptive content of "water", its Fregean sense. When, on the other hand, the chemist answers the question "What is water", he does not explicate the semantic structure of "water", but the chemical structure of water.

What is traditionally called a "nominal definition" is an explication whose object is a given term ("nomen"). By contrast, the object of a "real definition" is the thing ("res") denoted by that term. More precisely, the object of a nominal explication is the *meaning* of a given term, and the object of a real explication the *nature* or *essence* of the thing denoted by that term. The lexical explications of "water" to be found in dictionaries are examples of nominal explications, and the definition of water to be found in chemistry textbooks, "Water is H_2O", is an example of a real explication.

In the current debate about truth, the distinction between nominal and real explications is largely ignored. The reason probably is that, for most authors, this distinction is a "distinction without a

---

[10] I am presupposing here that "x is water" means "x is something having the same phenotype as water". In some contexts, "x is water" may mean "x is something having the same microphysical (chemical) structure as water (on earth)". For simplicity's sake, I shall ignore this problem.

difference". They identify the essence of a thing with the sense of the corresponding term. An exception is William Alston, who accounts for the distinction in an indirect way, by distinguishing between concepts and properties. In his seminal article "Truth: Concept and Property", he argues that "a property might have various features not reflected in our concept of that property" (2002, p. 12). Thus, the property of heat is revealed by physics to be the average kinetic energy of constituent molecules, even though our ordinary, pretheoretic concept of heat involves no such component. This concept consists of the set of features that any competent speaker associates with the word "heat", the "criteria" all speakers use when they apply that term. Obviously, ordinary speakers do not associate "heat" with the average kinetic energy of constituent molecules. For this reason, this feature does not belong to the ordinary concept of heat, though it may be a feature of the corresponding property.

According to Alston, the same applies to the concept and the property of truth: the property involves features that are not reflected in the concept. He argues that the minimalist account of truth, according to which the concept of truth is exhaustively explicated by the instances of Tarski's truth-scheme, is adequate when it is considered to be an explication of the concept of truth. The features of the property of truth that are not contained in the concept of truth and that are not captured by the minimalist account are made explicit by the correspondence theory of truth. Examples are the features that a true sentence is a sentence that corresponds to a fact and that true sentences are made true by facts. Just as the explication "Water is H_2O" correctly explicates the property of being water, but not the concept of water, so too the explication "A true sentence is a sentence that corresponds to a fact" explicates correctly the property of truth, but not the concept of truth.[11]

Frege's theory of sense and reference may also be regarded as an explication of the property of being true. According to it, the truth of a sentence depends on the sense and the reference of the words that occur in it. An elementary sentence of the form "a is F" is true

[11] This means, in Kantian terms, that this explication is not an analytical sentence, but a synthetic one. To arrive at it, it does not suffice to analyze concepts; it is necessary to analyze the structure of the world.

if and only if the object determined by the sense of "a" possesses the property determined by the sense of "is F". This explication is not meant as an explication of our concept of truth, which Frege considers to be primitive, but as an analysis of the semantic factors on which the truth of a sentence depends.

Elaborating Alston's distinction, we could say that concepts are constituents of thoughts and properties are constituents of facts. Just as two different senses expressed by proper names (such as the senses of "the evening star" and "the morning star") may refer to the same object, so too two different senses expressed by two one-place predicates (concepts) may refer to the same property, and two different thoughts may refer to the same fact.

A second possibility to account for the distinction between nominal and real definitions is based on the distinction between two senses of identity sentences such as "Water is H_2O", "Pain is C-fibre firing" and "Truth is correspondence with a fact". On the first reading, "is" means identity and on the second reading constitution.[12] Suppose, for instance, that properties are identical when the corresponding concepts are identical, that is, when the corresponding terms have the same Fregean meaning, the same criteria of application. In this case being water and being H_2O are different properties. Even in this case we can maintain that water is H_2O, in the sense that the property of being water is ontologically constituted (or "realized") by the property of being H_2O. Similarly, we may defend the thesis that truth is the correspondence with a fact, even when we admit that truth and correspondence with a fact are different properties.

On this approach, a real explication of a given property aims to make its ontological structure explicit – how the property is constituted. A nominal explication, on the other hand, seeks to make its semantic structure of the explicit – what the features are of which it is composed.

The distinction between nominal and real explications is important for at least three reasons. First, it shows that conceptual explications that are commonly considered to be incompatible might actually be complementary. Thus, if Alston is right, the minimalist

---

[12] This paragraph is inspired by the distinctions made in Brink 1989, pp. 156-163.

account of truth can be made compatible with the correspondence theory when the first is construed as an explication of the concept of truth and the second as an explication of the property of truth (in his sense of "concept" and "property").

Second, the distinction is important for the evaluation of Moorean arguments against some projects of conceptual explication.[13] To see this, consider the following example from metaethics. According to Moore's argument, a definition of morality in terms of "natural" terms like "A morally right action is an action maximizing human welfare" cannot be materially adequate, because the definiens and the definiendum have the same meaning. His argument is that sentence-pairs like "A morally right action is a morally right action" and "A morally right action is an action that maximizes human welfare" have different meanings: whereas the first sentence is analytic and trivial, the second is synthetic and informative.

This argument implies that the sentence "A morally right action is an action that maximizes human welfare" cannot be considered as an adequate explication of the *meaning* of "is a morally right action". But this does not rule out that it correctly explains the *reference* of "is a morally right action", what moral rightness is and in what it consists. The argument implies, in other words, that the project of giving a nominal explication of morality must be abandoned, but it does not imply that the project of giving a real explication of morality must also be abandoned.

Thirdly, the distinction between nominal and real explications is important for the discussion of deflationist arguments against philosophical theories of truth and non-cognitivist arguments against normative ethics. It is commonly assumed that a radically deflationist account of truth like the redundancy theory has devastating implications for those accounts which consider truth to be a property. For, if sentence pairs like "Snow is white" and "The thought that snow is white is true" express the same meaning, then there is no property of truth.[14] This consequence seems to imply that the endeavour of philosophical theories of truth to uncover the nature of the property of truth is doomed to failure right from the start. Similarly, the

[13] For more details, see Brink 1989, Chapter 6.

[14] This consequence is drawn, for instance, in Brandom 1998.

non-cognitivist accounts of morality, according to which "is morally right" is not used to ascribe a property, but to indicate the performance of an expressive or directive speech act, seems to undermine the possibility of normative ethics.

In fact, however, the consequences of the deflationist accounts of truth and the non-cognitivist accounts of morality are less dramatic, because the meanings of terms in natural language are relevant only for the adequacy of nominal explications, and not also of real explications. Suppose, for instance, that the redundancy account of truth is correct. In this case, there is no property to which the term "true" refers. But this does not imply that the property analyzed by the correspondence theory does not exist. If the correspondence theorist considers his account of truth as a real explication, he can argue that he is not interested in the semantics of "true", but only in the property to which he refers in his theory by means of "true". He may admit that, in natural language, "true" does not refer to any property, and to take this into account, he may replace this word in his theory by a similar one, say "frue", as Tarski suggested in a parallel case.[15] Natural language does not decide which properties exist. Similarly, the non-cognitivist account of morality does not imply that there are no ethical properties, but only, that these properties are not denoted by the corresponding terms of naturally language.

The distinction between intensional and extensional explications also refers to the objects of explications. In the case of intensional explications, the object is the meaning of the explicandum, and in the case of extensional explications, its extension. An example of an extensional explication is Tarski's definition of truth; its goal is to define the set of true sentences of the language of the class calculus.[16] His theorem of the undefinability of truth does not say that the meaning of the truth-predicate or the features of the property denoted by it cannot be made explicit, but only that in a semantically closed language it is impossible to construe in a consistent way an open sentence whose extension is the set of true sentences of that language.

We have seen that an important criterion of classification for

[15] Cf. Tarski 1944, p. 356.

[16] Cf. the definition of truth in Tarski 1933.

conceptual explications concerns the type of the explicandum. The explicandum can be either the meaning of a given term, or the property to which it refers, or the extension of this property. A second important criterion of classification concerns the purposes for which explications serve. In this regard, we must distinguish between descriptive, revisionary, reductive and stipulative explications.

4 Descriptive explications

A major part of the explications to be found in the history of philosophy are devoted to the task to make the content of a given concept explicit and to distinguish it from other concepts with which it is easily confused. Such explications may be called "descriptive" explications. The larger philosophical programs into which they are embedded pursue either a propaedeutic or a therapeutic goal. The descriptive explications can accordingly be subdivided into "propaedeutic" and "therapeutic" ones.

4.1 Propaedeutic explications

The context in which propaedeutic explications are typically made are the programs of the "metaphilosophical" disciplines such as meta-ethics, meta-logic, meta-metaphysics, meta-mathematics and so on. Their purpose is to prepare the construction of the corresponding philosophical discipline by making its central concepts clear and distinct.

Kant's explication of moral goodness is an example of an explication of this type. He distinguishes between good and evil on the one hand, and well-being and ill-being on the other, and then goes on to show that these pairs of concepts, which are easily confused, are independent of each other: it is not a contradiction that an evil man is doing well. According to Kant, this conceptual distinction is basic for the correct conception of ethics. An ethical theory that is concerned with well and ill-being is not an ethical theory in the proper sense, but a mere theory of prudence telling us how to act in order to have a pleasant life. An ethical theory in the proper sense tells us,

by contrast, what the difference is between good and evil actions.[17]

A second example is Frege's distinction between the concepts of being true and being taken to be true in his critique of the idealistic (or "psychologistic") foundation of logic.[18] These concepts, which are identified by the idealist logicians of his time, are independent according to Frege. What is taken to be true need not therefore be true. And it is not a contradiction that something which is taken to be true is actually false, he argues. In order to clarify the nature and the proper task of logic, it is essential, in his view, to distinguish clearly between these concepts. Thus, to discover the laws of truth is a task of logic, while to discover the laws of being taken to be true is not a task of logic, but of psychology.

The classical method to make the content of a given concept explicit is the decomposition of a concept into its conceptual constituents. Frege compares this method with the chemist's procedure to decompose a given substance into its molecular and atomic constituents. Just as the chemist analyses the chemical structure of substances, so too the philosopher analyses the logical structure of concepts. To display the logical structure of a given concept by means of an explication, the explicans must be syntactically more complex than the explicandum. Normally, the explicandum is syntactically simple and the explicans is composed of several predicates that are connected by logical signs. The meanings of the predicates occurring in the explicans are the conceptual constituents of which the explicandum is composed, and the manner in which the predicates are connected by logical signs represents the logical structure of the explicandum.

We may call explications whose method consists in the decomposition of concepts "analytic" explications. Obviously, this type of explication is applicable only to complex concepts. In order to make the content of an atomic (simple) concept explicit, we need an alternative method such as, for instance, the "holistic" method to determine the place of a given concept in our conceptual scheme. In this case the concept is not decomposed, but its conceptual connections with other concepts are made explicit. This procedure may be circular, i.e., it allows for the mutual explication of concepts. Thus,

[17] See Kant, 1788, p. 70.

[18] See the preface of Frege 1893.

Frege explicates the concept of judgment in terms of the concept of truth and vice versa: to judge is to acknowledge the truth of a given thought, and truth is the standard that we use to evaluate judgments in science.

If Donald Davidson is right, the attempt to decompose philosophical concepts into more fundamental concepts is a misguided project, because these concepts already belong to the most fundamental concepts we have. But this does not mean that these concepts cannot be explicated at all. In order, for instance, to explicate the concept of truth, which is considered by him to be primitive, we can adopt the holistic method to make its connections with other concepts like belief and linguistic meaning explicit.[19]

Quine also takes a critical stance on the conceptual program of traditional philosophy, but for another reason. In his view, it is impossible to distinguish objectively between "meaning and theory", between "analytic" and "synthetic sentences" and hence between conceptual analyses and empirical hypotheses.[20] Consider, for instance, the "folk theory of truth", which consists of the beliefs about truth shared by all competent speakers. This theory probably includes the beliefs that a sentence is either true or false, and that true sentences correspond to the facts. The problem posed by Quine's critique is that we do not have any method at our disposal that allows us to decide whether such a belief is analytic or synthetic. Since a conceptual analysis consists of analytic sentences, we consequently do not know how to decide whether an explication like "A true sentence corresponds to the facts" is to be considered as a conceptual explication or not. Applying the Moorean test, we might argue that this sentence is synthetic, because it is informative, in contrast to the tautological sentence "A true sentence is true". But this argument is not really conclusive because analytic sentences may also be informative. This is the case, for instance, when the components or the logical structure of the concept to be analysed are not obvious. Any explication revealing "hidden meanings" or "hidden semantic structure" is informative in some sense.

[19] Cf. Davidson 1996.

[20] Cf. Quine 1951 and Quine 1960, §§ 39 and 53.

4.2 Therapeutic explications

The second species of descriptive explications consists of the "therapeutic" explications, as we might call them alluding to Wittgenstein. These explications are also designed to make the content of concepts clear and distinct. However, in contrast to propaedeutic explications, their purpose is not to prepare the construction of a scientific discipline, but to show, on the contrary, that the endeavour to construct such a discipline is based on a linguistic misunderstanding.

An example of this type of explication is the minimalist account of truth advocated by Paul Horwich.[21] It falls into two parts. The first consists of an explication of the content and function of the word "true" according to which truth is a thoroughly formal, trivial and hence philosophically uninteresting concept that does not have any explanatory function in theories, but serves only certain technical needs that could also be achieved by quantifying over sentential variables. From this explication the consequence is drawn, in the second part of the account, that we do not really need a philosophical theory of truth.

Further examples of therapeutic explications are the non-cognitivist accounts of morality. Their purpose is to show that the endeavour to construct a moral theory is based on the linguistic misunderstanding that the predicate "is morally right" and its cognates are used to ascribe a property.

5 Revisionary explications

The context in which revisionary explications are made is the revision of theories. Their purpose is not to make the content of given concepts explicit, but to transform defective concepts into scientifically respectable ones. The defective concepts cannot simply be ignored because the play an important explanatory or expressive role in science. But they cannot be introduced directly into science because they are somehow defective. The aim of a revisionary explication is to revise the concept in such a way that, firstly, its explanatory or expressive function is preserved, and secondly, its defects are

---

[21] Cf. Horwich 1990.

eliminated.

Tarski's definition of truth is an example of this type of explication. According to him, the concept of truth is, on the one hand, indispensable for the construction of semantics and kindred disciplines, but, on the other hand, its direct introduction into science is illegitimate because of its inconsistency. The main purpose of his definition of truth is to transform the concept of truth into a concept that is both consistent and fruitful, that is, can be used to define the main concepts of logic and semantics like, for instance, the concept of logical consequence.[22]

The anti-realist explications of truth such as Dummett's verification theory are also of the revisionary type. They proceed from the assumption that the ordinary, realist concept of truth is scientifically not acceptable because it is potentially "recognition transcendent". If the truth of a sentence consists in its correspondence with reality, then there might be true sentences whose truth cannot be recognized, as, for instance, some undecidable sentences of set theory. To overcome this defect, the ordinary concept of truth must be replaced by a similar but epistemologically more adequate concept that can equally well "do the work" of the original concept, that is, that can be used to explicate the main concepts of logic and semantics (like logical consequence and linguistic meaning). The aim of the anti-realist explications of truth is to construct such a concept.

This example already shows that the realist- and anti-realist conceptions of truth are not necessarily incompatible. We may, for instance, accept the correspondence theory as a descriptive explication of truth, and the verificationist theory as a revisionary one.

In contrast to descriptive explications, a revisionary explication need not preserve the original meaning of the explicandum. To be materially adequate, it suffices that the meanings of the explicandum and the explanans are similar. To make this more precise, we need to explain what the conditions are for the explicandum and the explicans to be sufficiently similar with regard to their meanings.

To answer this question, some authors have proposed criteria re-

---

[22] The classical example of a revisionary explication is Kuratowski's definition of the concept of ordered pair in terms of the concept of set, which is considered by Quine as a "philosophical paradigm" (cf. Quine 1960, § 53).

ferring to the extension of the explicandum and the explanans. Carnap, for instance, adopted in the *Aufbau* the criterion of extensional identity. Goodman proposed the more liberal criterion of extensional homomorphism, and Quine the even more liberal criterion of extensional homomorphism.[23] These proposals ignore, however, what is essential for revisionary explications, namely, the preservation of the explanatory or expressive function of the explicandum. Hence, a revisionary explication should be considered as materially adequate if and only if the explicans can be used to perform this function.

Suppose, to illustrate this point, that the explicandum, say, the concept of truth, serves exclusively to refer to the members of its extension (true sentences) in a given theory. In this case the explicans must have, in order to fulfil this function, the same extension as the explicandum. If, on the other hand, the explicans is used also to refer to the concept of truth, then the explicans and the explicandum must have the same meaning.

The concept of number is a special case; its function is to count things. To fulfil this function, it suffices that the explicandum and the explicans are extensionally isomorphic. The reason is that every progression can be used to count things.

There are, however, cases in which even this liberal criterion is too narrow. The extension of the ordinary concept of existence is the set of existing objects. When this concept is explicated, as usual, as a concept of second order, the extension of the explicandum and the extension of the explicans do not even have any common member. The explicans does not apply to any object of which the explicandum is typically predicated. Nevertheless, this explication is materially adequate, at least when it is considered as a revisionary one. The defect of the explicandum is in this case that the truth-conditions of negative existence-statements like "Pegasus does not exist" are paradoxical, when existence is considered to be a concept of first order.[24]

These examples show clearly that the criteria of material adequacy for revisionary explications must be construed in a more flexible and differentiated way. Our criterion that a revisionary explication

[23] Cf.Carnap 1928, §§ 43-45, Goodman 1978, pp. 99-102, Quine 1964 and Quine 1969, pp. 55.

[24] Cf. Quine 1948.

must preserve the function of the explicandum takes this into account. Examples of explications that are not materially adequate on this criterion are the traditional explications of truth from the point of view of the disquotation theory. For, these explications are based on the assumption that the truth-predicate serves to ascribe a property. But, according to the disquotation theory, the truth-predicate cannot be used to this end, because it has a very special function, namely, to disquote quoted sentences.

A second shortcoming of the current debate about the adequacy of revisionary explications is that the criteria of adequacy discussed are incomplete. Thus, the adequacy of such explications depends also on the adequacy of the norms of scientific respectability that the explicans is supposed to satisfy. Some examples may illustrate this point.

The supreme norm in Frege's *Grundgesetze* for the methodologically sound introduction of new expressions into science proper is this: we have to "explain each expression with respect to its reference completely" (Frege 1903, §§ 57, 65). According to this norm, the use of an expression in science proper is legitimate only when its referential indeterminacy has been removed. This restriction is motivated by Frege's quest of ensuring the validity of the logical laws in rigorous science, in particular the validity of *tertium non datur.*[25]

A second norm due to Frege is the demand for a proper criterion of identity for abstract singular terms. "If we are to use the symbol *a* to signify an object", he writes, "we must have a criterion for deciding in all cases whether *b* is the same as *a*" (1884, § 62). This norm, which was later adopted by Quine, has come to be known as the norm "No Entity without Identity".[26]

Obviously, a revisionary explication that aims to adapt the explicandum to these norms is adequate only insofar as these norms are adequate. With regard to Frege's norm of referential determinacy, this is highly questionable, however, because it is impossible to determine the reference of all terms completely.

Secondly, the adequacy of a revisionary explication depends also on the adequacy of the method of explication it employs. This "meth-

[25] Cf. Frege 1903, §§ 55 ff.

[26] For a reconstruction of this norm, see Greimann 2000.

odological adequacy" of an explication depends, in the first place, on the norms of scientific respectability that are adopted. Suppose, for instance, that we want to transform the inconsistent ordinary concept of truth into a consistent concept. To ensure the consistency of the explicans, we must replace the explicandum by concepts whose consistency is beyond question. To this end, we may, for instance, follow Tarski's strategy to reduce the concept of truth to physical concepts.[27] Consequently, an explication seeking to transform truth into a consistent concept is methodologically adequate only if the method employed achieves a reduction (or "elimination") of the explicandum. The explicit definition of the explicandum is an example of a method satisfying this constraint, and the axiomatic characterization of the explicandum is an example of a method that does not satisfy it.

Thirdly, the adequacy of a revisionary explications depends also on the adequacy of the order of explication on which it is based. Such an order tells us which kind of concepts is to be explicated in terms of which other kind of concepts. Its task is to ensure that the scientifically defective concepts are explicated in terms of more respectable concepts, and not vice versa. Frege's explication of the concept of number in terms of logical concepts, for instance, is based on an order of explication according to which logical concepts are scientifically more respectable than arithmetical ones. To explicate logical concepts in terms of arithmetical ones would be a case of *clarum per obscurum*, from his point of view.

Finally, the adequacy of a revisionary explication depends also on its form, its structure. To be formally adequate, the explication must include all measures that are necessary to explicate the explicandum in a satisfactory way. Moreover, these measures must be taken in the correct order. The first step must be to determine the explicandum of the explication. Secondly, it must be shown that the explicandum is defective. Thirdly, the function to be preserved must be made explicit. Fourthly, a list of norms of scientific respectability must be set up and justified. Fifthly, an order of explication must be established that fixes the range of the possible explanantia. Sixthly, a method of explication must be selected by means of which the explicandum

---

[27] For more details, see Etchemendy 1988.

can be adapted to the norms of scientific respectability. Seventh, this method must be applied to the explicandum. Eighthly, it must be shown that the explicans satisfies the norms of scientific respectability.[28] As far as I can see, there is only one explication to be found in the literature that is formally correct in this extensive sense, namely, Tarski's definition of truth, which can therefore be considered as a paradigm of a revisionary explication.

6 Reductive explications

There are many programs in philosophy who are not concerned with the explication of a single concept, but with the explication of whole families of concepts. Examples of this type of explication are the reductive programs in the philosophy of mind, the philosophy of science and the philosophy of mathematics, like (analytic) behaviourism, phenomenalism, physicalism, and logicism. These programs do not aim to describe the structure of a single concept, but the structure of our overall conceptual scheme. According to (analytic) physicalism, for instance, our conceptual scheme has the hierarchical structure of a building whose fundament consists of the physical concepts.[29] All other families of concepts such as the psychological, sociological and semantic concepts occur at higher levels, and they can be reduced to the fundamental concepts via chains of explicit definitions.

The main purpose of a reductive explication of a given concept is neither to make its content explicit nor to transform it into a more respectable concept, but to show that it is in principle *superfluous*, in the sense that it can always be substituted by the explicans. An explication of this type may be considered as a rule for the translation of a formulation of a given theory that contains the explicandum into a formulation of the same theory that contains the explicans, but not the explicandum. Such a rule shows how the explicandum can be "eliminated", by showing how every sentence containing the

---

[28] In Greimann 2007, the formal adequacy of explications is described in more detail.

[29] There are weaker versions of physicalism that do not involve this thesis. They are not relevant in the current context.

explicandum can be translated into a sentence that does not contain it. In the case of logicism, for instance, the source language is the arithmetical language and the target language the language of logic.

To show that the explicandum is superfluous, it suffices to show that the function it performs in the theory can be performed equally well by the explicans. The criterion of material adequacy for reductive explications is therefore identical with criterion for revisionary explication: the function of the explicandum must be preserved.

The redundancy theory of truth can be considered as a reductive explication of the concept of truth. As already indicated by its name, the core of this theory is the doctrine that the word "true" is superfluous. The rule of translation it suggests for the translation of sentences in which "true" occurs is this: a sentence of the form *It is true that p* or *"p" is true* is to be translated simply as *p*.[30]

Tarski's definition of satisfaction and truth in terms of physical concepts (including set-theoretical ones) is a second example of a reductive explication. Unfortunately, this explication is not materially adequate because it does not preserve the intended function of the explicandum. For, the explicans is supposed to play an important explanatory function in the foundation of semantics, viz. the definition of the concept of logical consequence and kindred notions. This presupposes that the explicans can be used to explain the semantic truth-conditions of sentences. Actually, however, it cannot be used to this end, because the explication implies that a semantic explanation like "The sentence "Snow is white" is true if and only if snow is white" says the same as the tautological explanation "Snow is white if and only if snow is white".[31]

A second criterion of adequacy for reductive explications is that the translation rule must be general. In particular, the translation rule should be applicable to "new cases". To see this, suppose that the source language consists of exactly the following four sentences:

---

[30] To translate the quantified occurrences of "true", as in "All beliefs of Plato are true", we need to quantify over sentential variables. Thus, we can translate "All beliefs of Plato are true" as "For all p: if Plato believes that p, then p".

[31] For a more thorough discussion of this problem, see Etchemendy 1988.

Snow is white,
Snow is black,
"Snow is white" is true,
"Snow is black" is true".

In this case, we may translate the sentences containing "true" into a target language that does not contain this word by means of the following list-like explication:

x is true if and only if x = "Snow is white" and snow is white, or x = "Snow is black" and snow is black.

Though this explication does enable us to eliminate the truth-predicate, it is unsatisfactory because it does not show us how to eliminate the truth-predicate when it is applied to new sentences that we may introduce into the source language.[32] To overcome this difficulty, we need a more general rule of translation that accounts for such new cases.

7 Stipulative explications

Roughly speaking, a descriptive explication informs us about what the explicandum means in natural language, and a revisionary explication tells us what the explicandum should mean in science. The characteristic of a stipulative explication is that it informs us about how the author or speaker wishes the explicandum to be understood when she or he uses it.

Suppose, for instance, that an author wants to explicate the classical, Aristotelian concept of truth. In order to make clear that his explication refers to this concept of truth and not to another, he may stipulate that by "truth" he means "truth in the classical, Aristotelian sense". This stipulation is an example of a stipulative explication. A second example is Tarski's proposal to denote the classical, Aristotelian concept of truth by the artificial term "frue", if the majority of the theoreticians of truth should decide that the word "true" should be reserved to denote a different concept of truth.

Generally speaking, the purpose of a stipulative explication is to

---

[32] This criterion is suggested by Hartry Field's criticism of Tarski's definition of truth in Field 1972.

fix the intended meaning of the explicandum in order to use it as a technical term. The explications of this type divide into two species.

7.1 Explications of laziness

The main purpose of an explication of laziness is to abbreviate a longer expression. Thus, the stipulation to understand by "truth" the classical, Aristotelian concept of truth allows us to abbreviate the longer expression "truth in the classical, Aristotelian sense" by "truth".

From a theoretical point of view, explications of laziness are superfluous. They are relevant only for the linguistic quality of a theory formulation. This quality depends, for instance, on the readability and the elegance of the formulation. The criteria of adequacy for explications of laziness refer correspondingly to pragmatic and aesthetic qualities such as brevity and elegance of expression.

7.2 Metaphorical explications

The second species of stipulative explications consists of the "metaphorical" ones. Their purpose is to attach a new meaning to the explicandum that cannot be expressed by means of the expressions that are available in the language. In contrast to explications of laziness, metaphorical explications extend the expressive power of the language in an essential way.

An example of this type is Frege's explication of the concept of a "truth-value". According to it, the truth-value of a sentence is the circumstance of its truth or its falsity.[33] Its purpose is to inform the reader about what Frege wishes to understand by the "truth-value" of a given sentence or thought.

Syntactically, "the truth-value of x" is a functional expression. Applied to a proper name, it forms a more complex proper name. Semantically, it attributes to a given argument its truth-value, just as "the colour of x" attributes to a given argument its colour and "the temperature of x" attributes to a given argument its temperature. In natural language, we can express that the colour of snow is

[33] Cf. Frege 1892.

whiteness, but we cannot express that the truth-value of "Snow is white" is truth. The problem is that there is no expression of natural language that expresses the intended meaning of "truth-value". The purpose of a metaphorical explication is to fill such gaps. It does not abbreviate a longer expression, but provides a new meaning for an old expression.

Since the new meanings cannot be expressed by means of the expressions that are already at our disposal, it is strictly speaking impossible to fix the intended meaning of the explicandum. The only resources we have to explain what a truth-value is supposed to be are the expressions of natural language.[34]

For this reason, the success of a metaphorical explication depends on the imagination and the empathy of the hearer or reader. He or she must in a certain sense *guess* the intended meaning of the explicandum. The method we commonly use to overcome this difficulty is the metaphorical use of natural language terms. Thus, the designation of the truth-values as "truth-values" is supposed to give hints to the reader that enable him to guess or to infer in some sense the intended meaning of the explicandum. This metaphorical method is also used in the formation of concepts like "logical decomposition", "concept of second order", "analytic explication", "ontological commitment", "hypothetical imperative", and so on.

Since the majority of the technical terms used in philosophy are introduced in this way, the metaphorical explication is perhaps the most important type of explication in philosophy. Only very few concepts are formed by Quine's method of elimination (or reduction).[35] For this reason, the current debate about explication should pay more attention to this type.

Although metaphorical explications are based on stipulations, they are not arbitrary. In contrast to explications of laziness, a metaphorical explication can be successful only when the metaphorical meaning of the explicandum can somehow be derived from its literal meaning. Consequently, an explication of this type is materially ade-

[34] The application of ostensive definition and kindred methods is impossible in this case, because truth-values are abstract objects.

[35] In Greimann 2009, I have tried to show that Quine's ontological reductions by means of contextual definitions do not really work.

quate only when the intended meaning of the explicandum is similar or analogous to its literal meaning. Otherwise the explication would not contain any hints that the hearer or speaker needs in order to guess what the intended meaning is supposed to be.

Since the success of a metaphorical explication depends strongly on the imagination, the empathy and the scientific socialization of the hearer or reader, its material adequacy must be relativized to linguistic communities. An explication that is materially adequate for one community, say, the readers of Frege, may be inadequate for another community, say, the readers of Heidegger.

8 The adequacy of the suggested typology

I have tried to determine the main types of conceptual explications. The main criteria of classification I used to this end refer to the objects and the purposes of explications. With regard to the objects, we must distinguish between nominal and real explications, and between intensional and extensional ones. With regard to their purposes, conceptual explications divide into five main groups: descriptive, revisionary, reductive and stipulative definitions. The descriptive explications can be subdivided into propaedeutic and therapeutic ones, and the stipulative explications into explications of laziness and metaphorical explications. In order to refine this typology, we could classify explications with regard to their methods and their form.

The adequacy of the suggested typology is not a question of its truth or falsity, but of its utility. It may be useful mainly with regard to the following three ends.

First, it enables us to evaluate the adequacy of a given explication in a more sophisticated way. Consider, for instance, Tarski's definition of truth. It does not consist of a single explication, but of various different explications that serve different purposes and have different criteria of adequacy. The first step of Tarski's explication consists of a stipulative explication by means of which the intended meaning of the definiendum is fixed. With regard to its purpose, the definition involves a descriptive explication of truth that is supposed to "capture" this meaning, the classical concept of truth. At the same time, the definition contains a revisionary explication that is supposed to replace the ordinary concept of truth by more respectable concepts.

With regard to its object, the definition must be considered as an extensional explication whose object is the set of true sentences of the object language. Nevertheless, it can also be considered as an intensional explication designed to explain the meaning of "true" from the classical point of view. Finally, the definition can also be considered as a real definition telling us on which language-world relations the property of truth depends. In order to evaluate Tarski's definition adequately, we must take into account that it contains various explications of very different types and with different criteria of adequacy. It might turn out that some of these explications are satisfactory and some not.

Second, our typology may contribute to the task tackled by Kirkham to resolve some of the confusions to be found in the current debate about truth. The typology shows, for instance, that the realist and the anti-realist theories of truth must not necessarily be seen as competitors, because they may also be regarded as complementary explications with different goals and different criteria of adequacy.

Finally, and most importantly, the typology may also contribute to the project envisaged by Carnap to construct an explicit theory of conceptual explications in which the rules for the correct explication of concepts are discussed.[36] To discuss these rules, we need to know the criteria of adequacy for explications. And to explain these criteria in an adequate way, we must distinguish between different types of conceptual explications, because the criteria of adequacy of a conceptual explication depend on its type.[37]

Dirk Greimann
Fluminense Federal University

References

Alston, William. 2002. Truth: Concept and Property. In *What is Truth?*, ed. by Richard Schantz. Berlin, New York: de Gruyter, pp. 11-26.

[36] In Greimann 2007, a system of rules for the correct revisionary explication of concepts is developed.

[37] This paper is based on a talk that I delivered at the Universität Hamburg, the Universität Münster and the Universidade de São Paulo in 2010. I am indepted to Guido Imaguire for his suggestion to write a paper on this topic.

Brandom, R. 1998. *Making it Explicit.* Second edition. Cambridge/M.: Harvard University Press.

Brink, David. 1989. *Moral Realism and the Foundations of Ethics.* Cambridge/M.: Cambridge University Press.

Carnap, Rudolf. 1928. *Der logische Aufbau der Welt.* Berlin.

Carnap, Rudolf. 1950. *Logical Foundations of Probability.* Chicago.

Davidson, Donald. 1996. The Folly of Trying to Define Truth. *Journal of Philosophy* 93, 263–278.

Etchemendy, John. 1988. Tarski on Truth and Logical Consequence. *Journal of Symbolic Logic* 53, 51–79.

Field, Hartry. 1972. Tarski's Theory of Truth. *The Journal of Philosophy* 69, pp. 347-375.

Frege, Gottlob. 1884. *Die Grundlagen der Arithmetik. Eine logisch-mathematische Untersuchung über den Begriff der Zahl.* Breslau, reprint: Darmstadt, 1961.

Frege, Gottlob. 1892. Über Sinn und Bedeutung. In *Funktion, Begriff, Bedeutung.* Göttingen: Vandenhoeck und Ruprecht, 1980, pp. 40-65.

Frege, Gottlob. 1893. *Grundgesetze der Arithmetik. Begriffsschriftlich abgeleitet*, Vol. I. Jena, reprint: Darmstadt, 1962.

Frege, Gottlob. 1903. *Grundgesetze der Arithmetik. Begriffsschriftlich abgeleitet*, Vol. II, Jena, reprint: Darmstadt, 1962.

Frege, Gottlob. 1914. Logik in der Mathematik. In *Nachgelassene Schriften*, ed. by H. Hermes, F. Kambartel and F. Kaulbach. Hamburg: Felix Meiner, 1983, pp. 219-270.

Goodman, Nelson. 1978. *Ways of Worldmaking.* Indianapolis: Indiana.

Greimann, Dirk. 2009. Contextual Definition and Ontological Commitment. *Australasian Journal of Philosophy* 87, pp. 1-17.

Greimann, Dirk. 2007. Regeln für das korrekte Explizieren von Begriffen. *Zeitschrift für philosophische Forschung* 61, pp. 261-282.

Greimann, Dirk. 2000. No Entity without Identity: A Reductionist Dogma? *Grazer Philosophische Studien* 60, pp. 13–29.

Gupta, Anil. 2008. Definitions. *Stanford Encyclopedia of Philosophy*, online edition, http://plato.stanford.edu/entries/definitions.

Horwich, Paul. 1990. *Truth.* Oxford: Basil Blackwell.

Kant, Immanuel. 1788. *Kritik der praktischen Vernunft.* Hamburg: Meiner, 1990.

Kirkham, Richard. 1995. *Theories of Truth. A Critical Introduction.* Cambridge/M.: Cambridge University Press.

Quine, Willard van Orman. 1948. On What There Is. In *From a Logical Point of View.* Second edition. Cambridge, Mass.: Harvard University Press 1964, pp. 1–19.

Quine, Willard van Orman. 1951. Two Dogmas of Empiricism. In *From a Logical Point of View.* Second edition. Cambridge/Mass. Harvard University Press, 1964, pp. 20-24.

Quine, Willard van Orman. 1960. *Word and Object.* Cambridge/M.

Quine, Willard van Orman. 1964. Ontological Reductions and the World of Numbers. In *The Ways of Paradox and Other Essays.* Second edition. Cambridge/M., pp. 212–220.

Quine, Willard van Orman. 1969. Ontological Relativity. In *Ontological Relativity and Other Essays.* New York, pp. 26–68.

Tarski, Alfred. 1935. Der Wahrheitsbegriff in den formalisierten Sprachen. In *Logik-Texte*, ed. by K. Berka and L. Kreiser. Darmstadt, 1983, pp. 445–546.

Tarski, Alfred. 1944. The Semantic Conception of Truth and the Foundations of Semantics. *Philosophy and Phenomenological Research* 4, pp. 341-376.