Metaontological Studies relating to the Problem of Universals

Mika Oksanen

Academic dissertation to be publicly discussed, by due permission of the Faculty of Arts at the University of Helsinki in auditorium XII of the Main Building, on the 11th of January, 2014 at 10 o’clock.
Acknowledgements

I am deeply indebted to the supervisor of my dissertation, Professor Gabriel Sandu, and to the preliminary examiners of the dissertation, Professor Jussi Haukioja and Docent Markku Keinänen for their comments.

I also owe thanks to Docent emeritus S. Albert Kivinen, who supervised my licentiate thesis and whose lectures and seminars on ontology have inspired the trains of thought which eventuated in this dissertation. I wish to thank Professor Ilkka Niiniluoto, who examined my licentiate thesis and whose lectures and writings have also greatly influenced me. I also wish to thank many students of Docent Kivinen and other persons interested in ontology as graduate students or researchers in the University of Helsinki with whom I have often discussed ontological questions, including Dr. Anssi Korhonen, Dr. Heikki J. Koskinen, Dr. Tuomas Tahko and Mr. Janne Hiipakka and many others I cannot here name separately.

Work on this dissertation has been financially supported by Finnish Cultural Foundation, Emil Aaltonen Foundation and University of Helsinki, which is hereby gratefully acknowledged.

This dissertation has been written with the aid of the freeware \LaTeX\ typesetting program.

I owe perhaps the greatest thanks to my mother Ulla Oksanen for her financial and emotional support during my lengthy work on this dissertation.

I dedicate this work to the memory of my deceased father Arvo Oksanen.

Hyvinkää, December 2013

Mika Oksanen
Abstract

My dissertation deals with metaontology or metametaphysics. This is the subdiscipline of philosophy that is concerned with the investigation of metaphysical concepts, statements, theories and problems on the metalevel. It analyses the meaning of metaphysical statements and theories and discusses how they are to be justified. The name ”metaontology” is recently coined, but the task of metaontology is the same as Immanuel Kant already dealt with in his Critique of Pure Reason. As methods I use both historical research and logical (or rather semantical) analysis. In order to understand clearly what metaphysical terms or theories mean or should mean we must both look at how they have been characterized in the course of the history of philosophy and then analyse the meanings that have historically been given to them with the methods of modern formal semantics. Metaontological research would be worthless if it could not in the end be applied to solving some substantive ontological questions. In the end of my dissertation, therefore, I give arguments for a solution to the substantively ontological problem of universals, a form of realism about universals called promiscuous realism. To prepare the way for that argument, I argue that the metaontological considerations most relevant to the problem of universals are considerations concerning ontological commitment, as the American philosophers Quine and van Inwagen have argued, not those concerning truthmakers as such philosophers as the Australian realist D. M. Armstrong have argued or those concerning verification conditions as such philosophers as Michael Dummett have argued. To justify this conclusion, I go first through well-known objections to verificationism, and show that they apply also to current verificationist theories such as Dummett’s theory and Field’s deflationist theory of truth. In the process I also respond to opponents of metaphysics who try to show with the aid of verificationism or structuralism that metaphysical questions would be meaningless or illegitimate in some other way. Having justified the central role of ontological commitment, I try to develop a detailed theory of it. The core of my work is a rigorous formal development of a theory of ontological commitment. I construct it by combining Alonzo Church’s theory of ontological commitment with Tarski’s theory of truth.
Contents

1 Introduction ............................................. 5
   1.1 What are Metaontology and Metametaphysics and why are they needed? .................. 5
   1.2 A Roadmap of the Argument .................... 15

2 History of the Concept of Ontology .......... 19
   2.1 History of Metaphysics from Aristotle to Wolff ............ 19
      2.1.1 Aristotle’s different Definitions of Wisdom ........ 20
      2.1.2 The Division of Metaphysics and Birth of Ontology in
            the Beginning of the Modern Age .................. 25
      2.1.3 Recent History of the Concepts of Metaphysics and
            Ontology: Heinrich Scholz and the Phenomenological
            Tradition ............................................. 31
      2.1.4 Ontological Fundamentalness .................... 38
   2.2 Kant’s Metametaphysical Attacks on (Transcendent) Meta-
       physics ............................................. 48
      2.2.1 Four major Epistemological Positions and their metaon-
            tological Consequences .......................... 51

3 Verificationism and Other Metaontological Principles in Log-
   ical Positivists ........................................ 68
   3.1 Different Elements in Logical Positivism: Empiricism, Struc-
       turalism, Foundationalism, Coherentism and Syntacticism ... 68
      3.1.1 Empiricism and Verificationism in Logical Positivism . 68
      3.1.2 A Proposed Re-evaluation of Logical Positivism .... 71
      3.1.3 Structuralism and Structural Realism in Logical Pos-
            itivism and Outside It ............................ 76
      3.1.4 Coherentist and Conventionalist Elements in Logical
            Positivism .......................................... 86
      3.1.5 Syntacticist attacks on Metaphysics ................. 89
   3.2 Logical Positivists’ Arguments against Metaphysics and Their
       Problems ............................................ 95
   3.3 Problems with Verificationism in General .......... 111
Chapter 1

Introduction

1.1 What are Metaontology and Metametaphysics and why are they needed?

This dissertation is concerned with metaontological or metametaphysical questions relating to the Problem of Universals. Metaontology or metametaphysics is the subdiscipline of philosophy that is concerned with the investigation of metaphysical statements on the metalevel, with analysing the meaning of metaphysical statements and terms and discussing how metaphysical statements and theories are to be justified. In this it is similar to the somewhat better known and more popular discipline of metaethics, which is concerned with analysing the meaning of ethical statements and discussing the ways in which they can possibly be justified. We can call both metaethics and metaontology parts of metaphilosophy.

The name ”meta-ontology” or ”metaontology” (both spellings are used in the literature) is the common name of the discipline\(^1\). The word ”meta-

\(^1\)The origins of this common name are a bit obscure; as is suggested by C. Daniel Dolson in [Dol06, page 13], it may have been first introduced by Jack Guendling in the rather curious article [Gue53] which Guendling himself according to Dolson later described as whimsical. Guendling defines meta-ontology in [Gue53, page 219] as that empirical branch of semiotic which studies symbols which, as related in the construction of theoretical systematics of objective relations among things, are completely general.

The definition is a bit obscure and has objectionable features. It should not be part of the definition of meta-ontology that it is empirical, for it is a difficult question which methods are to be used in meta-ontology and rationalistic theories of meta-ontology are certainly possible. It is surely plausible that if any statements at all are analytic, then some statements of meta-ontology are, and traditional analytic philosophy would therefore
ontology” was memorably used and popularized in its modern sense by Peter van Inwagen in [vI98]. Inwagen calls the question ”what are we asking when we ask what is there?” the meta-ontological question and any answer to it a meta-ontology\(^2\). It is natural to expand the term ”meta-ontology” also to the discipline that seeks such answers (especially since the word ”ontology” has already earlier been used both of theories and a discipline in a similar way). The word ”metametaphysics” is also occasionally used in the literature, though it is very ugly; a whole book with this as its title, namely [CMW09], has recently appeared. Unfortunately, most articles in this book concentrate on discussing the possibility of metaphysics rather than actually seeking to constructively provide methods for metaphysical research\(^3\).

Since an adequate treatment of metaontology requires a lot of time and care, I must devote most of this dissertation to considering metaontological questions. However, metaontological research would be worthless if it could not in the end be applied to solving some substantive ontological questions (in the possibly weak sense of the word in which a typical ontological question can be substantive). At the end, therefore, I will give a preliminary argument for a solution to the substantively ontological problem of universals, a form of realism about universals called promiscuous realism. It seems to me that the argument I will give will lend very strong support for the claim that some kind of realism concerning universals is true. If properly understood, this claim will turn out to be almost obviously correct.

While the problem of universals is not metaontological but substantive, it is one of the least substantive problems of substantive metaphysics. While it is then not wholly trivial, it is more trivial than other metaphysical problems. proclaim them a priori. However, the emphasis on complete generality is surely on the right track, as we will see later in Section 5.6 of this work.

\(^2\) This definition of the term is simpler and more comprehensible than Guendling’s, though it may be a bit too narrow, as it concentrates on the meaning of ontological questions and neglects the question of how they are justified, which can also be called metaontological.

\(^3\) The words ”ontology” and ”metaphysics” are usually employed as synonyms, and therefore the words ”metaontology” and ”metametaphysics” are naturally also used as synonyms. However, I will show later in Section 2 that a difference is yet often made between the two words ”ontology” and ”metaphysics”, so that the former is a part of the later or a more fundamental discipline on which the later is based. This might make it advisable to also make a difference between ”metaontology” and ”metametaphysics”, so that the former would be a part of the later or a more fundamental discipline on which the latter would be founded. However, the distinction between the two disciplines would in any case be rather small, so this is a subtlety that may not be of much practical importance.
I think that it is also independent of other, more substantive problems. I will argue that the realism concerning universals I will defend can be combined equally well with different solutions to other metaphysical problems, such as the problem of the relation of mind and matter, which is perhaps the most substantive (and most difficult) of all metaphysical problems. I will argue that realism concerning universals can be combined equally well with physicalism, dualism, idealism and neutral monism.

I think that attacking ontological questions directly, in a naive way, is not the most promising way of trying to make real progress on them. Ever since Immanuel Kant introduced what he called critical philosophy - and even earlier, when Rene Descartes raised the cogito, an epistemological principle as the starting point of philosophy - such a naive way of proceeding in ontology has been held in suspicion. Kant’s *Critique of Pure Reason* can be held to have been an early work of metametaphysics (though it concentrated exclusively on the epistemological question of the justification of metaphysical statements, assuming uncritically that the meaning of such statements was sufficiently clear). Kant addressed the same theme more succinctly in *Prolegomena To Any Future Metaphysics That Will Be Able to Come Forward As Science* (translated into English in [Kan85]). Metaontology or metametaphysics can be said to be a prolegomenon to metaphysics in Kant’s sense. Metametaphysics is thus far from a completely novel discipline, though it has had few practitioners until recently. Most proponents of modern analytical philosophy share this old suspicion. However, several current ontologists, especially outside and on the peripheries of analytical philosophy, have defied this general opinion and continued trying to attack ontological questions directly. I do not think that they give good enough arguments for reverting to such an antiquated procedure. They are with good reason disgusted by the extreme subjectivism into which the critical method has often led; however, there is no reason to think that a critical method would necessarily lead to subjectivism- Such modern meta-ontologists as van Inwagen have not been led to any subjectivist or sceptical conclusions about ontology. We can suspect that only a bad use of the method does so. On the other hand the naive approach to ontological questions has led to unsolvable disagreements. Therefore I will develop my ontological theory with the aid of preliminary metaontological or metametaphysical considerations.

In reaction to the negative attitude towards ontology and metaphysics
that has prevailed in analytical philosophy, many ontologists have gone to
the other extreme and defended the primacy of metaphysics or ontology
(which is sometimes - e. g. by C. B. Martin in [MH99] and John Heil-
 called the ontological turn or sometimes the speculative turn). Some of
them, such as Michael Devitt who speaks about the priority of metaphysics e.
ge. in [Dev01a], are naturalistic metaphysicians; however, many of them (e. g.
Barry Smith in [Smi] or Tuomas E. Tahko in [Tah08]) are Neo-Aristotelians
and think that they are returning to the Aristotelian approach. Tahko says
in the abstract of his thesis:

The metaphysics which I support could be called Aristotelian as
opposed to Kantian: metaphysics is the first philosophy and the
basis of all other philosophical and scientific inquiry.

Neo-Aristotelians often say that when Descartes placed epistemology in
the center of philosophy instead of ontology, things started going downhill; Barry
Smith’s video lecture available on the Internet [Smi] is a very good example
of this very widespread attitude. However, there are reasons to think that
such Neo-Aristotelians have partly misunderstood the approach of Aristotle
or are at least unclear about in what sense metaphysics can be primary.

Aristotle distinguished two kinds of priority in Posterior Analytics (72a);
he said that there is a difference between what is prior and better known (or
intelligible) in the order of being and what is prior and better known (intelli-
gible) to man. The former kind of priority might be called metaphysical
or ontological priority, while the later might be called epistemic or episte-
omological priority. A discipline can be said to be prior to another in one of

4These classes are not mutually exclusive; some Neo-Aristotelian metaphysicians are
also naturalistic metaphysicians. However, this is not always the case. There exist both
more empiristic and more rationalistic interpretations of Aristotle’s philosophy, and both
more materialistic and more dualistic interpretations of it. While the more empiristic and
materialistic interpretations of Aristotelianism are clearly compatible with naturalism,
the more rationalistic and dualistic interpretations are not. E. g. Smith’s and Tahko’s
interpretation of Aristotelianism is so rationalistic that it is incompatible with the kinds of
naturalism professed by Devitt or Ladyman, which combine physicalism with empiricism.
Some who defend the primacy of ontology are Heideggerians, with a very idiosyncratic
conception of what ontology is, a conception that Heidegger indeed pretended to base
loosely on Aristotle’s conception, but developed in a way very different from that in which
most Neo-Aristotelians (or analytical metaphysicians) develop it.

5Exactly how this notion of metaphysical priority is to be understood is a difficult
question, and what entities (if any) are metaphysically prior to others, i. e. fundamental
is already a very substantive metaphysical question. I will touch these questions briefly
later in many sections of this dissertation, mostly in Section 2.1.4.
these senses if the truths it tries to find out are prior to the truths the other
seeks in that sense. I think that more subtle philosophical analysis must
actually distinguish several different kinds of epistemological priority and
several different kinds of metaphysical priority: however, at this point the
distinction between these two broad kinds of priority is enough. I am afraid
that philosophers who speak of the primacy of metaphysics or ontology often
confuse even these two kinds of priority.

How are these priorities related? Aristotle tells us in Metaphysics (1029b3-
5) (see [Ari33, pages 318,319]) that

\[...\overset{\text{learning is always acquired in this way, by advancing through}}{\text{what is less intelligible by nature to what is more so.}}\]

This makes quite clear at least that what is ontologically primary cannot in
all cases be the same as what is epistemologically primary, since we need to
advance through one to the other. It also suggests that these two kinds of
priority are reverses of each other. However, this seems to be too simple;
other sayings of Aristotle imply they cannot be so in all cases; for instance,
Aristotle suggests in Metaphysics 1028a that substance is primary in all
senses, both by definition and in knowledge. Many later metaphysicians,
not just Aristotelians but also rationalists, have followed him in this and
interpreted it so that substances are both metaphysically and epistemologi-
cally primary.

It seems to me that the truths of metaphysics are indeed according to
the Aristotelian view prior and better known \textit{in the order of being} than the
truths of other disciplines, but they are generally very far from being \textit{prior
and better known to man}. A discipline can be said to prior and better known
in one sense if and only if its truths are prior and better known in that
sense. Therefore metaphysics would (unsurprisingly!) be metaphysically
prior to other sciences; however, in this case it could not be prior to them
epistemically. In this case it cannot be the basis of philosophical inquiry as
Tahko claims, at least if this expression is understood in the most natural
sense, since the basis of inquiry is most naturally understood as that from
which the inquirer starts, i. e. that from which he advances\footnote{Metaphysics can indeed be the basis of inquiry in the sense that the holding (not the
knowledge) of its truths makes inquiry possible in the first place, but this would be a
rather unnatural way to use the expression.}. This seems
to me to be an intuitively plausible view about the relationship of these
disciplines; it is also supported by naturalistic metaphysics, since in special sciences also two kinds of priority can be distinguished. We can speak about an entity having a more fundamental role in a scientific theory than another, and a naturalistic metaphysician would naturally take metaphysical priority to be some kind of generalization of this kind of primacy in a scientific theory. While middle-sized physical objects are epistemologically prior, yet science takes microscopic particles which we cannot directly observe to play a more fundamental role in a scientific theory.

Willem de Jong has in many articles, e.g. in [dJB10] and in [Jon95], discussed the Classical Model of Science that derives from Aristotle’s *Posterior Analytics* aka *Analytica Posteriora*. Descartes was according to de Jong [dJB10, page 187] also a follower of this tradition and the second milestone of this tradition after Aristotle’s *Posterior Analytics* was the *Logic of Port-Royal* (translated in [ArnCL]) written mainly by the Cartersian Antoine Arnauld and relying in many respects on Pascal and Descartes⁷. Jong has stressed the importance of this distinction, which earlier commentators on Aristotle’s methodology such as Heinrich Scholz and Evert W. Beth in his view [dJB10, page 197] missed, leading them to a too harsh criticism of the model (which however is still outdated in many ways, as its claim that the principles of science must be known infallibly is just not sustainable.). As de Jong shows, this distinction corresponds to a distinction between two correlative methods, the method of synthesis and the method of analysis.

There is some dispute among historians of philosophy about the origin and meaning of the very word ”metaphysics”. Many (perhaps the most) say that the etymology of the word has nothing to do with its meaning and that it originated just because Aristotle’s book *Metaphysics* dealing with the discipline was placed after the book *Physics* in the edition of Aristotle’s works (perhaps by Andronicus of Rhodes, who according to some not too reliable sources named it on this basis). However, some say that it was intended to signify that metaphysics comes after physics when studying philosophy (which then meant the sciences in general). On the other hand, Aristotle also calls metaphysics first philosophy (*πρωτη ϕιλοσοφια* in Greek, *prima philosophia* in Latin), which suggests that it is prior to physics and other

---

⁷ Arnauld gives Descartes’s proof for the immortality of the soul as one example of analysis; see [ArnCL, page 307]. This argument of course begins from Descartes’s notorious attempt to doubt everything that can be doubted, which Barry Smith thinks leads to scepticism.
special sciences. However, even if we think that the name “metaphysics”
did connote the supposed place of metaphysics as a discipline in the system
of sciences, these two views about the relationship of metaphysics to physics
and other special sciences can be reconciled if we distinguish the two kinds
of priority. We can then say that metaphysical truths are prior and better
known in the order of nature than the truths of special sciences like physics,
and therefore metaphysics is first philosophy in the order of being, but the
truths of special sciences like physics are prior and better known to man,
and therefore metaphysics comes after physics when studying the sciences.

Nevertheless, even the truths of special sciences are not best known to
man, but better know to man than them is the pre-scientific (often common-
sensical) knowledge that precedes all of them. Before having any knowledge
of electrons or black holes, we have pre-scientific knowledge of men and rocks
and trees. Even prior to this we can argue (as I will later in this work argue)
to be the experiences of individual persons, including at least their percep-
tion of material objects and probably far more, such as inner perception
and mathematical intuition. Therefore we cannot start from metaphysics or

---

8When Quine famously objected to first philosophy in the name of naturalism, he was
not attacking first philosophy in the order of being but first philosophy in the order of
knowledge. In other words, it was traditional foundationalist epistemology that he was
attacking, not metaphysics or ontology as he is sometimes misunderstood to have done.
Even so, the theory Quine himself presented in e. g. [Qui93] seems to be foundationalist
in a weak sense, as Quine accepts the existence of observation sentences as expressing
the foundations of science. Therefore Quine must either be acknowledged to be inconsis-
tent or he must be interpreted so that it was only some specific powerful forms of
foundationalism that he was attacking. In [Qui90], Quine admitted in response to the
article [Haa90] by Susan Haack that his theory is foundationalist, though it combines
foundationalist elements with coherentist ones; Quine even said that his theory might be
called foundherentist, if the word coined by Haack were not so ugly. Donald Davidson
described Quine’s (as well as Dummett’s) view as foundationalist in [Dav86, page 312] and
attacked it from a coherentist point of view, claiming that it led to skepticism. Davidson
held in [Dav86, page 313] that we should give up the distinction between observation sen-
tences and other sentences, and this seems to me to be indeed a consequence of complete
anti-foundationalism. Davidson indeed does not distinguish verificationism clearly from
weaker forms of foundationalism. Yet it seems that what the later Quine mostly objected
to is the kind of foundationalism according to which the foundations of knowledge are not
intersubjective, such as the theory that claims the foundations of knowledge to concern
sense-data; he is at least sometimes implicitly willing to accept the kind of foundationalism
according to which the foundations of knowledge concern physical objects. When Devitt,
who follows Quine, attacks first philosophy (e. g. in [DS87, page 225-226]), it seems that
he is primarily thinking of rationalistic foundationalism as his foe, as he contrasts first
philosophy with philosophy naturalized, which according to him says that philosophy is
not an a priori discipline. Devitt’s attack on first philosophy is then consistent with the
acceptance of empiricist foundationalism. I will discuss different kinds of foundationalism
and their differences at more length later in Section 3.1.1 of this work.
even from physics, but we have to start from the pre-scientific knowledge or from experiences preceding even it and proceed from this by the method of analysis\(^9\) or resolution or inquiry (zetesis) to the truths and concepts that are better known in the order of being, and then use these principles to derive the truths of the special sciences and semantics and logic themselves by the method of synthesis or composition. This will confirm that our analysis has been successful and also provide us with more truths of the lower level sciences. In fact the principles of logic, semantics, epistemology and methodology, which Aristotle treats of in his *Organon*\(^{10}\) govern the whole processes of analysis and synthesis and are therefore in one way prior to and better known to man than all the truths that are known with the aid of these methods. Unfortunately, Aristotle himself did not develop the methodological disciplines of analysis (with the exception of logic) very thoroughly, but skipped swiftly to the later stage of synthesis, which may account for the errors that have later been discovered in his theories. The later development of science has shown that many of Aristotle’s doctrines that Aristotle himself thought to be certainly known, such as his physical doctrines of the four elements or of the heavenly spheres, were in fact false, so the processes of analysis and synthesis could not provide such infallible knowledge as Aristotle thought. Even Aristotle’s logic, though it remained unsurpassed for much longer than his physics, has been found defective by

---

\(^9\)The method of analysis as the phrase is here used must not be confused with analytical knowledge nor the method of composition with synthetic knowledge, as these words are commonly used in modern philosophy. There are indeed interesting (although obscure and controversial) connections between these two notions, especially in Kant’s thought as de Jong shows in [Jon95]. However, all the same the method of analysis may lead to synthetic knowledge as well as analytic, since it may start from truths known (at least relatively) immediately on the basis of experience, in which case it will lead to synthetic knowledge, or it may start by examining only the concepts abstracted from experience, in which case it will lead to knowledge of conceptual relationships, which is commonly called analytical knowledge.

\(^{10}\)Aristotle did not yet have any names for semantics or epistemology but at least his work *De Interpretatione* treats of subjects that we would today consider as semantical and his work *Posterior Analytics* treats of subjects that we would today call epistemological and methodological, while *Prior Analytics* treats of logic. As for *Categories*, it is controversial whether it treats of ontology, semantics or even grammar. If Aristotle’s philosophy was better than earlier and competing ones, this is to a great extent due to the fact that he had a better organon than others, including the first highly developed theory of logic in the theory of syllogistic inference in *Prior Analytics*. However, this is not generally true of present day Neo-Aristotelians. Neo-Aristotelians who talk about the primacy of metaphysics often either use Aristotle’s outdated organon or do not have any organon at all, which leaves their metaphysics built on a foundation of sand.
modern logicians like Frege, Russell and Whitehead, even though some Neo-Scholastics (such as Henry Babcock Veatch with his intentional logic) have still during the past century tried to defend it against modern logic. There is surely no reason to suppose that Aristotle’s more purely philosophical doctrines such as his metaphysical doctrines would be any more generally true than his physical doctrines, as some of the more extreme Neo-Aristotelians (especially Neo-Scholastics, though even secular Neo-Aristotelians are occasionally guilty) do. This is especially so since many if not most metaphysical principles are posterior to and less well known to man than physical ones, and therefore the falsity of Aristotelian physics implies that any metaphysical conclusions based on that physics are utterly unjustified once the falsity of Aristotle’s physics is recognized. As an especially egregious instance, it is obvious that Aristotle’s argument for the existence of God as the Prime Mover (or probably rather of several gods as prime movers) in Book XII of Metaphysics (1073b-1074a) was based on the current physical theory of the motion of the stars (and the spheres which he thought to carry them) developed by such astronomers as Eudoxus and more generally on his dynamics (which Galileo showed to be false) and is hence pretty much worthless today (though it it can of course not be ruled out before detailed investigations in the philosophy of religion that it might be possible to develop variations of it that might be based on or at least compatible with modern physics). However, the same considerations are likely to extend far wider to Aristotle’s metaphysical principles.

Nevertheless, they do not apply to all of Aristotle’s metaphysical claims; as I will later show in more detail in this work (in Subsection 2.1), Aristotle thought that the principle of non-contradiction (formulated in one way) was a metaphysical statement, and it is not at all plausible to claim that our knowledge of the principle of non-contradiction would depend on physics or other special sciences. It may not be at first sight plausible to claim that

11This is not wholly uncontroversial. As I will discuss in more detail later in Section 2.2.1 of this work, extreme empiricists would no doubt be willing to make even this claim. Holistic empiricists such as Quine or Devitt would claim that even such logical laws as the principle of non-contradiction would be in principle revisable in the light of physical knowledge; however, most even of them would think that the development of physics has in actuality given no reason to revise it. Some dialetheists like Routley would claim that the principle should be revised, and would appeal to empiricism in favour of its revisability. However, it seems to me that the real main reasons why they want to revise it are not based on empirical sciences, but derive from such reasons as set-theoretical and semantical paradoxes, which in their view are best solved by accepting that they validly
the principle of non-contradiction would be epistemically primary either, at least not in all senses of that expression; even being able to state that principle requires the possession of concepts that are acquired only with difficulty in a high state of civilization. However, it can be argued that once a person is capable of clearly formulating and understanding the principle, its truth is immediately obvious to him, so in another, more important sense it can be claimed to be epistemically primary. If this is correct, it is a plausible example of a rare principle that is both ontologically and epistemically primary. Nevertheless, it appears unlikely that very much of Aristotle’s metaphysics is in this way independent from his science. Also the fact that Aristotle’s metaphysics contains so very different kinds of claims raises doubts whether it is really a unitary discipline, or whether it should be divided into subdisciplines as we will later see Christian Wolff proposed should be done.

In this circumstance a sharper scrutiny of the methods of analysis and a concentration on epistemology and methodology is surely called for. Descartes’s epistemological theorising was in fact motivated by the desire to recover the method of analysis used by ancients such as Aristotle and to develop it further; Descartes’s revolt against the medieval interpretations of Aristotle was one of the many causes that led to the rise of modern science. There is certainly much room for disagreement about how well he succeeded in his task and about the ultimate correctness of his positive epistemological theories, especially with regard to his representationalism about perception. However, the epistemological problems (such as the problems posed by illusions and hallucinations) he posed are logical developments of the problems already present in ancient philosophy (e. g. in the thought of the Sceptics) and no serious philosopher today can avoid engagement with them. Modern analytical philosophy, as can be seen in its very name, can also be understood as an attempt to understand the process of analysis better. We

imply contradictions. However, the paradoxes are surely not in any significant way based on special sciences. However, Aristotle would also have thought that the Principle of Excluded Middle was part of metaphysics, and it has often been claimed that discoveries in physics affect that principle; e. g. Putnam has claimed in [Put68] that Quantum Mechanics bids us to reject that principle. However, given that there are many competing interpretations of quantum mechanics, an interpretation that demands us to modify logical principles is surely one of the least attractive. Of course, conceptual analysis in the sense of the word used in analytical philosophy is only a small though extremely important part of the method of analysis in the broadest sense, which also includes such methods as deduction, induction and abduction (if abduc-
must then apply the knowledge gained in these disciplines to metaphysical
questions, and doing this is the task of metaontology or metametaphysics.

1.2 A Roadmap of the Argument

Metaontological theories can be divided into semantical and epistemological
ones. It is very important to distinguish these kinds of problems from each
other; as I will argue, many influential philosophers such as logical positivists
have confused them, and this has led them to confusions which still persist
in today’s philosophy. In this dissertation I concentrate on the semantical
side of metaontology. I will often refer to epistemological problems, but will
not take any position with regard to most of them. However, I must take
a position in the basic epistemological debate between foundationalism and
coherentism. I will argue that coherentism is utterly untenable and there-
fore we should either accept foundationalism (though a very weak version
of foundationalism is sufficient) or (if possible) seek out a third option, such
as e. g. Susan Haack’s foundherentism is supposed to be. Nevertheless, my
arguments are designed so that they will work equally well in very different
epistemological frameworks, between which I do not at least yet have rea-
sons to decide; for example, they are designed to be compatible both with
(at least many kinds of) representationalism and direct realism, with both
more empiricistic and more rationalistic theories, and with externalism as
well as internalism. I will not be able to even discuss the debate between
internalists and externalists. However, I have to discuss the debate between
extreme empiricists and their opponents, since realism about universals has
traditionally been associated with rationalism and this debate has therefore
been connected with the Problem of Universals. I do so in Section 2.2.1
of this work, but only to show that it is so difficult that we should not at
present rely too much on any conclusions with regard to it in metaontology,
and is not as closely connected with the Problem of Universals as is often
supposed.

I will consider the meaning of ontological claims and the epistemological

value of different ways of supporting or attacking such claims before trying to actually justify any substantive ontological claim. Many metaontological considerations are common to several ontological problems; for instance such commonly used ontological concepts as ontological commitment, truthmaking, reduction, supervenience, ontological dependence etc. can (if they are meaningful and non-trivially applicable to anything, which can of course in many cases be questioned, as I will show in the case of truthmaking) be applied to many, perhaps any substantive ontological problems. Therefore the theories of such concepts are independent of any specific ontological problem. However, there are also those metaontological considerations that are specific to some single problem of substantive ontology. Since the ultimate aim of this dissertation is finding an (at least preliminary) solution to the problem of universals I will discuss mostly the metaontological problems associated with the problem of universals\(^{13}\). I will attack the problem of universals indirectly by discussing what the problem is and how it could be solved. This involves studying how such general metaontological concepts as verification, ontological commitment and truthmaking can be applied to the special problem of universals. I will argue in Chapter 6 that - contrary to what has been often claimed recently - the metaontological considerations most relevant to the problem of universals (a substantive problem of ontology) are considerations concerning ontological commitment, not those concerning verification conditions or those concerning truth-makers. To justify this conclusion, I go first in Chapter 3 through well-known objections to verificationism, and show that they apply also to current verificationist theories such as Dummett’s theory and Field’s deflationist theory of truth. In the process I also respond to opponents of metaphysics who try to show with the aid of verificationism or structuralism that metaphysical questions would be meaningless or illegitimate in some other way. I then discuss in Chapter 4 the reasons for thinking that truths in general have truthmakers. I will come to the conclusion that the reasons generally given for these suppositions are insufficient, so that while I do not have any conclusive evidence

\(^{13}\)Another very central, perhaps even the most central, metaphysical problem is that of the relation of mind and matter. I will not try to give any, even preliminary, solution to it in this dissertation. However, I must occasionally refer to it as it is closely tied up with the most fundamental epistemological and semantical questions. I will also use it as an example of a metaphysical theory, since it rather uncontroversially is one, and discuss what it is that makes it a metaphysical theory.
that axioms such as the truthmaker axiom are false, yet we should suspend judgement on them unless or until better reasons for them are found. Naturally this has implications for metaontology and ontology and metaphysics as a whole; I will argue that ontology and metaphysics are not wholly explanatory disciplines but in large part (though not wholly) descriptive.

I will argue for the inseparability of systematic and historical research in many subdivisions of philosophy, at least at the current state of its development. Analytic philosophers have often been accused of a lack of historical awareness, and this accusation is often justified. However, there have also been many analytical philosophers who have fruitfully combined historical research with logical analysis, such as for example Wolfgang Künne, Nino P. Cocchiarella and Paulo Crivelli. I will base my work partly on their research and develop it further. The importance of historical research in philosophy has been stressed by such philosophers of science as Thomas Kuhn in [Kuh70]; however, accepting the importance of historical research by no means requires the adoption of such a relativistic standpoint as Kuhn arrived at.

A historical perspective is especially important for metametaphysics, for metametaphysics tries to clarify what the problems of metaphysics are, and this requires knowledge of their historical development. Lack of historical knowledge of how the problems that are currently discussed have developed often leads to a misunderstanding of the nature of such problems. I will argue that such philosophers as Patterson misunderstand what the correspondence theory of truth is and such philosophers as Gonzalo Rodriguez-Pereyra misunderstand what the problem of universals is because they are not sufficiently aware of the history of these theories and problems.

More generally, many current analytical philosophers remain influenced by doctrines of logical positivism such as verificationism, often without being aware of this, though such doctrines have long been refuted or strongly disconfirmed. I will argue that Field’s deflationary theory of truth is vitiated because it is to a great extent motivated by verificationism.

On the other hand there has also been overreaction to positivistic tendencies, such as the claims about the primacy of ontology I already mentioned. Another possible instance of such overreaction is the overemphasis on explanation. Such philosophers as Pierre Duhem (see [Duh91]) had thought that natural science did not explain, but only metaphysics did. Logical posi-
tivists had accepted Duhem’s view on science, but thought that metaphysics was impossible, which resulted in a view that had no place for explanation. Probably because of overreaction to this very implausible view, many current philosophers seem to find explanation central in both natural science and philosophy. My scientific realists hold inference to the best explanation or abduction to be central to natural science, as it gives us knowledge of theoretical entities. In philosophy it is held to be essential to the correspondence theory of truth that truth is explanatory. Many problems of philosophy such as the problem of universals are held to be problems of how to explain something.

I would rather say that what is central to natural science is extrapolation rather than explanation. I will argue that it is induction rather than abduction that gives us knowledge of theoretical entities. What makes natural science practically important is that it can help us to predict the future and prepare for it, though for theoretical purposes other kinds of extrapolation such as retrodiction are equally important. I do not deny that natural science also helps us to explain why many things are as they are, but I would view this as rather a byproduct of extrapolation rather than central to natural science. Philosophy, however, seldom explains anything in the way natural science does. I will argue as an instance of this in Chapter 4 that explanation is not essential to the correspondence theory of truth and in Chapter 6 that it is not central to the problem of universals either.
Chapter 2

History of the Concept of Ontology

2.1 History of Metaphysics from Aristotle to Wolff

In order to know what can be and should be meant by the words "metaphysics" and "ontology" we must look at how they have been used during the history of philosophy. I must therefore provide a history of the notions associated with these words. It will only be a history of the very notions of the disciplines and their definitions, not of substantive metaphysical and ontological theories. Indeed, it would be impossible in the bounds of this work to give any kind of general history of ontology that would not be so simplified as to be misleading; libraries are full of short sketches of the history of metaphysics that are so truncated as to be all but useless. However, it is possible to give a sketch of the very notions of metaphysics and ontology that is comprehensive enough for my purposes.

Ontology and metaphysics are often taken to be identical, and even when this is not the case they are taken to be closely related\(^1\), so we must begin with the concept of metaphysics, since it is historically older and the concept of ontology has developed from it.

\(^1\)Many philosophers assume that there must be some difference between them, but since they often are ignorant of the history of the development of those words and the associated concepts, they seem to just choose the differentiating factor arbitrarily, often arriving at weird distinctions which have nothing to do with the history of these concepts. All too often the resulting distinction between ontology and metaphysics is a distinction without difference, so that the resulting two concepts turn out to be equivalent.
2.1.1 Aristotle’s different Definitions of Wisdom

One of Aristotle’s definitions of metaphysics (which Aristotle called Wisdom\(^2\) or first philosophy – \(σοφία\)) in book Γ of Metaphysics (1003a) was that it is the science of being as being \((ον \ η \ ον)\), i. e. the only science that dealt with all that is or exists, i. e. deals with what is common to all existing entities. The truths of metaphysics are then truths that are more general than the truths of any other sciences, absolutely general. I will later in Section 5.6 examine more thoroughly what this means and whether there can be such truths. Metaphysics is therefore the only science that had to take account in some sense of everything that there is. Obviously if its aims are to be realistically attainable this must be in some rather weak sense, rather in the sense of providing room for everything in its theory - e. g. giving for every entity a category under it falls (and then studying the interconnections of these categories) - than in the sense of deriving all entities and all of the laws governing them from some first principles.

The definition of metaphysics as the science of being as being is closely connected to the later (post-Aristotelian) definition of metaphysics as the science concerned with the world as a whole, i. e. the definition of metaphysics as cosmology - in a sense different from physical cosmology, which is a part of physics. This definition of metaphysics has been common among analytical philosophers. E. g. Bertrand Russell characterized metaphysics in [Rus18b, page 1] as the attempt to conceive the world as a whole by means of thought. Similarly Moore characterized the main problem of Metaphysics in [Moo53, page 25] as the general description of the Universe. The definitions are close since the world or universe (in the widest senses of the words, as distinct from the sense in which they have stood for a planet or a galaxy) is supposed to be a being which contains all beings (with the possible exception that theistic metaphysicians often want the world to only contain all created beings). Therefore the definitions are equivalent if such a being exists. However, if metaphysics is defined as the science of being as being

\(^2\)Though it serves to muddle matters further, we must note in the interest of historical accuracy that Aristotle seems to have used the word “wisdom” sometimes for philosophy as a doctrine (the outcome of philosophical reflection, as is indeed etymologically natural) as a whole and sometimes for metaphysics as a proper part of philosophy. Indeed Aristotle, and many of the later philosophers that followed him, including his commentators, seem sometimes to have used the word “philosophy” to stand for metaphysics alone and at other times to stand for the sum of all sciences or scholarly disciplines.
this definition does not presuppose that a complex entity such as the world
exists (though it allows it to exist), so Aristotle’s older definition is better
since it has fewer presuppositions.3

Matters are complicated by the fact that Aristotle gave other defini-
tions or characterizations for wisdom. Aristotle also characterized it in
Metaphysics (see [Ari33, 932 a, page 9]) as concerned with the primary
(prota) causes (aitia) and principles (arkhas) (a characterization of meta-
physics which Giovanni Reale in [RC80] calls aetiology4; this must of course
be distinguished from the more common use of the word in medicine to refer
to the study of the causes of diseases.).

Aristotle also characterized metaphysics (in 1064a 34-35; see [Ari35,
pages 86-87]) as the science which deals with that which exists separately
(χωριστον) and is immovable (ακινητον). Exactly what Aristotle means
by separation (or separability, as Donald Morrison has in [Mor85] shown
it might also be translated) is highly obscure and controversial5. In any

3Wolff already distinguished cosmology, which he held (e. g. in [Wol40, §77, page 35])
to be part of physics, and can be seen as a primitive version of physical cosmology, from
general cosmology (see [Wol40, §78, page 36]), which he held (e. g. in [Wol40, §79, page
36]) to be a part of metaphysics.

4Reale also distinguishes a definition of metaphysics as ousiology, i. e. theory of sub-
stances; however, this is rather one way of understanding the definition of metaphysics as
ontology rather than an apparent rival to it. Even if it were decided that ontology is a the-
ory concerning being as being, that still leaves open the question whether metaphysics is
only a theory of entities that exist in the sense that substances do, or also of entities which
exist in the sense that entities from other categories do. Of course, here we suppose that
this classification of entities into categories is correct and that entities of these categories
exist in a different sense, which of course a modern philosopher must question when he is
doing metaphysics and not just history of metaphysics. A definition of metaphysics cannot
exclude process metaphysics, according to which substances either do not exist at all or
at least are not fundamental, as a metaphysical theory. E. g. Whitehead’s process philo-
sophical theory would be commonly called a metaphysical theory. Admittedly Marxists
do often say such things as that metaphysics is a way of thinking which thinks of things
in abstraction from their change and development; however, this way of understanding
metaphysics does not correspond to the way the word has been used almost anywhere
outside Marxism. Indeed Marxists contrast metaphysics with dialectical thinking, which
they have taken from Hegel, whose system of dialectics is generally taken to be an extreme
example of metaphysics at its very worst.

5This is not an unimportant issue. Since Aristotle in arguing against Plato’s meta-
physics constantly opposes Plato’s view that the Ideas exist separately, this notion is vital
to the understanding of the difference between Aristotle and Plato and of Aristotle’s the-
ory of universals. The notion is also important for Aristotle’s theory of the agent intellect
in his De Anima, which is described as separate, and therefore for the controversy about
how naturalistic his philosophy of mind was. However, modern commentators have not
been able to come to any sort of agreement or clarity about exactly what this separation
means; the notion of separation might be understood in terms of independence (i. e. illu-
minated with the aid of modal concepts) or in mereological terms or simply as numerical
case it is rather clear that separation means separation from matter, so this
definition implies that metaphysics is a science which deals with substances
which are in some sense immaterial. Aristotle himself immediately goes on
to say that if there exists something which is separate and immovable, it
is divine (θειόν, theion), so Aristotle understood metaphysics on the basis
of this characterization also as theology; however, since it is unclear what
separateness means it is also unclear why Aristotle thought that something
separate would have to be divine.

It is obscure and controversial how these other definitions or character-
izations are related to the definition of metaphysics as the science of being
qua being and how far they agree with it. Different (groups of) comment-
tators and followers of Aristotle have stressed different definitions through
the ages. The ancient Greek commentators mostly viewed metaphysics as
theology, the study of the highest, immaterial substances; of the Arabic
commentators Ibn Sīnā aka Avicenna stressed the definition of metaphysics
as the science of being qua being\(^6\), while Ibn Rushd aka Averroes stressed
the definition of metaphysics as theology; the medieval Scholastic commen-
tators mostly stressed the definition of metaphysics as the science of being
as being. It must be noted that there is good textual evidence that Kant
himself understood the dogmatic metaphysics he criticized as natural theol-
ogy; he characterizes metaphysics in the Critique of Pure Reason (B7 A3)
as the science that, with all its preliminaries, has for it especial object the
solution of problems of pure reasons God, freedom of the will and immor-
tality. What Kant thinks to be mere premiminaries to genuine metaphysics
would by Wolff have been understood as an independent discipline of ontol-
ogy, a subdiscipline of metaphysics. This gives us reasons to suspect that
Kant’s critique of traditional metaphysics is just irrelevant against those

\(^6\) This can be seen from an English translation of one of his works on meta-
physics \([Avi05]\); in \([Avi05, page 3]\) Avicenna argues that God’s existence cannot be the
subject matter of metaphysics, though it is one of the things investigated by this science.
In \([Avi05, page 9]\) Avicenna summarizes his discussion by saying that
the existent inasmuch as it is an existent is something common in all these
things and that it must be made the subject matter of this art for the reasons
we have stated.

The phrase "this art" refers to metaphysics, as is clear from the preceding context. Inter-
estingly, Avicenna anticipated Wolff in saying (see \([Avi05, page 11]\)) that the science
of metaphysics is necessarily divided into parts, something that most exponents of the
theological interpretation of metaphysics would deny.
traditional metaphysicians (and their modern successors) that understood their discipline as the science of being as being. The German Neo-Kantian philosopher Paul Natorp was perhaps the first at least in the modern age to see how problematic the diversity of Aristotle’s definitions was. Not only the various definitions Aristotle gives, but the actual discussions in different parts of Aristotle’s work seem to be concerned with different subjects;

The existence of these different conceptions has often been explained with the aid of different far-going historical hypotheses. Natorp thought that the definition of metaphysics as theology was a later Platonic interpolation. While this hypothesis has later been generally abandoned, the diversity has often been accounted for with the aid of developmental hypotheses concerning Aristotle’s thought. Werner Jaeger was the first to suppose (e.g. in [Jae12] and in [Jae34, page 218]) that the different definitions derive from different periods in Aristotle’s philosophical development; the conception of metaphysics as theology from the time when he was yet a faithful Platonist and the definition of metaphysics as the science of being qua being from a later time when he rebelled against Plato’s ideas. Such developmental hypotheses have been much criticized (see [RC80] for a discussion). It is not needful for my purposes in this work to discuss in detail whether such hypotheses are correct or not. This is fortunate, since the task might be impossible; it is likely that not enough data survives to reliably reconstruct the development of Aristotle’s thought. I will mention that the many apparent contradictions between many of Aristotle’s statements do strongly suggest that they derive from different periods of his thinking, but it is probably impossible to say which are earlier and which later; however, we certainly should not assume that the later views are necessarily better, so the relative chronology of the different views does not matter much. Fervent opponents of developmentalism like Wehrle in [Weh00, page 1] say that the developmentalists’ version of Aristotle at best strikes one as being a bit of a bungler and at worst a confused dabbler, changing his mind as regularly as Bertrand Russell but without the ability to keep straight just what he had thrown out and what he retained. However, just changing his mind surely does not make Aristotle a bungler (nor does it make Russell a bungler); all philosophers who keep working on their problems conscientiously eventually change their mind on important issues. Developmental hypotheses have been suggested for the work of almost all great philosophers. It can be proposed that it was the editors of Aristotle’s writings - Eudemus or Andronicus of Rhodes etc. - rather than Aristotle himself who could not keep straight what he had thrown out and what he had retained. In any case, other great philosophers have been accused of a similar inability even when it is known that they themselves have put their work together - e.g. Kant’s Critique of Pure Reason has been considered as a patchwork. Even current philosophers, who unlike Aristotle have computers and printers to help them, have great difficulty in keeping straight what they throw out and what they retain. So it is not on the face of it unlikely that Aristotle would have shared this inability. We certainly should not idolatrously view Aristotle as a superman who unlike all other philosophers never made a mistake and never had to revise his views. Nevertheless when it comes to the specific question of the definition of metaphysics and the division of the parts of philosophy, it may be possible to reconcile the different definitions without the hypothesis of a change of theory on Aristotle’s part, just given some background assumptions. However, while this may be interesting from a purely historical point of view, it is of no help for a contemporary metaphysician or ontologist who is trying to find out the boundaries of his discipline. Such assumptions are highly doubtful from a contemporary point of view, and surely cannot be taken for granted and therefore cannot be presupposed as part of the very definition of metaphysics, so a modern metaphysician must choose between the different definitions.
while Aristotle’s discussions of actuality and potentiality, form and matter, etc. are indeed relevant to his proof of the existence of God or Gods, yet they are independent of them in the sense that they could be accepted by persons that do not think the proof to be valid, and Aristotle himself uses the concepts developed in them outside theology, for instance in his physics.

It is easy to fail to see how profound the difficulty of interpreting Aristotle here is. It can indeed be argued quite plausibly, as it has often been argued\textsuperscript{8}, that the other definitions of metaphysics imply the definition of metaphysics as the science of being as being given some assumptions Aristotle made. It can be argued quite plausibly that in speaking about the most primary causes and principles of we are also speaking (even if only implicitly) about everything, for if say the Prime Mover is a cause (or even just the mover) of everything, then it is true of everything that it is caused (or moved) by the Prime mover. However, the definitions would only be equivalent if not only the other definitions implied this definition, but also this definition implied them. However, it is very hard indeed to argue with any plausibility that it would do so.

This can be clarified by noting that a discipline is individuated by the questions it examines, and a question can be understood as the set of its answers - as is frequently done in the logic of questions - i.e. the propositions or sentences or other truth-bearers whose truth it examines. Therefore two definitions identify the same discipline only if every proposition answering a question examined by the discipline identified by one of them answers a question examined by the other. All that the kind of argument we are examining would show is that any proposition answering a question asked by metaphysics as aetiology or theology also answers a question asked by metaphysics as the science of being as being; it does not show that any proposition answering a question asked by metaphysics as the science of being as being would answer a question asked by aetiology or theology. Therefore this line of argument would only show at most that metaphysics as aetiology or theology is a part of metaphysics as the science of being as being where.

\textsuperscript{8}E.g. Merlan argues in [Mer68, page 173]:

This uppermost sphere of being somehow "causes" all the other spheres and its elements are the elements of everything. Therefore, the true philosopher . . . deals with the elements of this uppermost sphere and thus with being. By implication, he therefore deals with being as it is present everywhere.
being, not that the disciplines identified by the definitions would be identical.

One could surely claim for example that every entity as such is a substance or an accident (which is a proposition whose truth is examined by ontology i.e. general metaphysics, whether a true proposition - as Aristotle assumed - or false - as process metaphysics assumes) without considering whether such propositions as that the Prime Mover exists (a proposition examined by aetiology and theology) are true even if it were true that every entity is caused by the Prime Mover. Of course if the first proposition were true then the Prime Mover would also be a substance or an accident. However, a physical being would equally well have to be a substance or an accident. Nevertheless, the applicability of the metaphysical principle to physical objects would not according to the Aristotelian conception of the division of sciences imply that metaphysics as the science of being as being would be part of physics. An Aristotelian would rather say that physics presupposes that principle but does not itself examine whether it is true (while a modern naturalistic metaphysician would rather say that physics implies that principle). Therefore neither would the applicability of the principle to the Prime Mover imply that metaphysics as the science of being as being would be a part of aetiology or theology. Rather it would, analogously to the case of physics, follow from this applicability and the way Aristotelians divide sciences that aetiology and theology would presuppose such principles of general metaphysics without examining their truth.

2.1.2 The Division of Metaphysics and Birth of Ontology in the Beginning of the Modern Age

Later, in the beginning of the modern age, when metaphysics was thought to extend beyond the limits of this definition of it as being qua being, ontology was defined as the most fundamental part of metaphysics, while the study of immaterial substances was delegated to the less fundamental science of natural theology or (if all immaterial substances were not thought to be divine) pneumatology, which was defined as the general science of spirits i.e. immaterial substances. Thus ontology came to correspond to metaphysics in the sense of the science of being qua being.

Most of the first philosophers who referred to their work as ”ontology” defined ontology very similarly to Aristotle’s definition of metaphysics
as the science of being as being. Indeed, the very word "ontology" comes from the Greek word for being (as did equivalents such as "ontosophy"), so the meaning of the word "ontology" is clearly bound up with the notion of being even more than the word "metaphysics". However, Leibniz’s

---

9The history of the origin of ontology as a new science is quite complex. Finding out the origin of the word "ontology" has turned out to be a work requiring lots of painstaking detection among historians of philosophy and historians of ideas, and may not be at an end. I will briefly describe this history and detection for the interested on the basis of the work of Raul Corazzon in [Cor11] at http://www.ontology.co/pdf/history.pdf and Leo Freuler and Jose Ferrater Mora in [Mor63], which I have checked from primary sources and emended so far as I have been able to expending a reasonable amount of work. It is not necessary to follow the following excursus carefully to understand the main points of this chapter, but those who want to assure themselves of the historical soundness of my view of ontology rather have to go into the details and if they do not truth my presentation check it themselves in Corazzon’s and Freuler’s and Mora’s works and in the (usually Latin) primary sources which they can find. It was long believed that the first use of the Latin form of the word, "ontologia", was in the Cartesian Scholastic Johannes Clauberg (1622-1665), then it was thought that it was found in Rudolf Göckel’s (aka Rudolf Goclenius’s, 1547-1628) Lexicon Philosophicum from 1613, where Goclenius in a marginal note defines ontology (written in Greek) as "philosophia de ente", philosophy concerning being (or beings). However, still newer research indicates that the origin of the word is in a book by Jacob Lorhard (aka Lorhardus); this was first thought to be his Theatrum Philosophicum from 1613, but it was later discovered that Lorhard had earlier used the word in an earlier version of the same book, Ogdoas Scholastica, from 1606 (which again was based on an earlier work, Clemens Timpler’s Metaphysicae Systema Methodicum, where however the new word does not yet appear). After this such philosophers as Johannes Clauberg used the word "ontologia", together with an alternative, "ontosophia" i.e. "ontosophy" - wisdom concerning being rather than doctrine concerning being. This was a time when many new terms for scholarly disciplines were coined, of which some came later into general use while others (noology, angelography, etc.) have been utterly forgotten; even the word "psychology" dates from this time. However, the coining of a new word and the discovery of a new discipline are not at all the same thing, and the earliest uses of the word "ontology" are philosophically quite uninteresting, even if they must be mentioned for the sake of historical accuracy. The first philosophers who used the word did not usually say much about what they considered ontology to be and when they said something they often identified ontology with metaphysics, as is generally done today, but nevertheless gave different definitions to the two words so that even if they thought them to denote the same science they yet differed in meaning, ontology being the science of being as being. However, soon a difference was made between the disciplines denoted by the two terms, though of course even then all who used the two words did not distinguish their meanings in the same way. The same distinction had of course already been done (e. g. by the Jesuit Benito Pereira (1535-1610), who spoke of metaphysics as a general science and natural theology as a special science) without using these exact terms - instead of ontology and metaphysics philosophers spoke of general metaphysics and special metaphysics, etc.

10Leibniz himself apparently followed Aristotle in using all of his definitions and taking them to be equivalent, at least if we can trust the interpretation of Ross in [Ros88]; Leibniz apparently characterized metaphysics both as the discipline concerned with those things that are common to every genus of beings - i. e. as the science of being qua being - and as the science of immaterial reality. While Wolff is often thought to be little more than a pedantic systematizer and popularizer of Leibniz, this shows that he yet had some originality in his division of the philosophical disciplines and other such methodological
follower and popularizer Christian Wolff was the most famous of the early philosophers to use the word. In \[\text{Wol}30, \S 1\] (see \[\text{Wol}05, \text{pages} 8-19\] for a modern German translation) he defines ontology as the science of Being in general or of Being as Being (scientia entis in genere, seu quatenus ens est).

This change is obviously connected with the secularization of metaphysics. If all three Aristotelian definitions of metaphysics are assumed to be equivalent (or indeed even if just one-way implication is assumed between them), this implies assuming that a materialistic world-view is necessarily false, since it implies that the same science deals with the most fundamental entities and immaterial entities, and therefore that the most fundamental entities are immaterial. However, the mere assumption that metaphysics deals with separated i. e. immaterial substances already contains very strong religious assumptions. The immaterial substances (such as Aristotle’s Intelligences or prime movers) have usually and naturally been understood as religiously significant entities, either gods (in polytheistic religions such as the Greek religion probably presupposed by Aristotle and his followers) or the one God or (if God is understood as being so high that he cannot even belong to any genus, as was commonly the case in philosophies affected by Neoplatonism) as angels (in monotheistic religions as in Judaism, Christianity and Islam, to which medieval and many later Aristotelians belonged). Therefore if reference to separate entities is taken to be part of the very definition of metaphysics then the existence of gods is presupposed by the possibility of metaphysics, and therefore religious presuppositions are connected with the very notion of ontology.

Of course, the definition can be weakened so that such presuppositions are eliminated, and metaphysics would be taken to deal with the question of whether separate entities exist and if so what they are. However, while such a weakened definition would avoid controversial presuppositions, it would yet not apply to all enquiry that has been commonly thought of as metaphysical in later philosophy; even philosophers who are atheists and materialists can disagree sharply about many questions which would commonly be taken to be metaphysical (e. g. are the fundamental entities material substances or material processes, etc.).

However, the definition of metaphysics or ontology as the science of being as being does not have any such obvious religious presuppositions. There has points; and as we shall see, these original ideas of Wolff proved influential and fruitful.
indeed been one way to interpret that definition that would cause it to have such presuppositions. Neoplatonistic philosophers reified being into Being Itself and went on to deify it. This kind of interpretation can be found for example in the commentary of the late pagan Neoplatonist Syrianus on Aristotle’s metaphysics, as seen for example in the following quotation (typically sublime or pompous - according to your taste - in its style) from its translation in [Syr08, page 90]:

For either there is is something which is, and is nothing else (such as intellect, or soul, or the heavens, or the world), uniquely existing as what is, as Being itself (autoon), clearly more worthy than all else, itself not needing to be a world, or the heavens, or soul, whereas all the rest receive their being from it.

While this is presented as a disjunction, it is clear from the context that Syrianus accepted the first disjunct (and in fact probably both disjuncts). Syrianus understood this Being Itself as a god, though not the highest god in his complicated system, since Unity itself or the One (identified by him with the Good) was higher than even it. Most medieval scholastics, though Aristotelians rather than Platonists, such as Thomists, yet followed Neoplatonists in this, since they had to accept as authoritative the writings of church fathers (like Augustine) who were Christian Neo-Platonists and were influenced by pagan Neo-Platonists; they of course had to identify this Being itself with the highest since the sole God of their theology. Martin Heidegger, the founder of existentialism, followed this line of thought in his most famous work, translated in [Hei62], as he also held that ontology had to consider Being (Sein) itself and not beings, though he opposed what he called ontotheology and so was not willing to identify this Being itself with any god as the Neo-Platonists and scholastics did (though the way in which Heidegger speaks of Being has a very mystical and reverent sound, such as is typically associated with talk regarding something divine, even if not always personal). There are some modern interpreters who think that this interpretation or a very similar one is correct; e. g. Philip Merlan in [Mer68].

However, this theory is on the face of it highly counterintuitive and was partly made just so that the two notions of metaphysics would coincide. Furthermore, the theory is more likely to come from Platonism and Pythagoreanism (as Syrianus partly admits) than to have been part of Aris-
totle’s original intentions, even though medieval Aristotelians who mixed Neoplatonic ideas with Aristotelian ideas adopted it. Indeed, in Metaphysics 1060a38-1060b6 (see [Ari35, pages 60-63]) Aristotle seems to have explicitly argued against this kind of theory, arguing that Being itself or Unity itself cannot be principles (as Syrianus later supposed). Of course, this does not imply that the Neoplatonic theory of Being itself and Unity itself would not be metaphysical theory, as it clearly is a metaphysical theory, though not an Aristotelian metaphysical theory. However, it indicates that unless Aristotle was blatantly inconsistent, he cannot have intended that his definition of metaphysics as the science of being qua being would be read so that it would have implies such a theory since he himself rejected it. Therefore such a theory cannot belong to metaphysics on the basis of its definition.

This kind of interpretation seems to me to misunderstand how the expressions ”being qua being” or ”being as being” and indeed how expressions like ”qua” and ”so far as” are used. I think S. Marc Cohen puts the error in the Neo-Platonic interpretation well in [Coh12]:

So Aristotle’s study does not concern some recondite subject matter known as being qua being. Rather it is a study of being, or better, of beings - of things that can be said to be - that studies them in a particular way: as beings, in so far as they are beings.

In any case, whatever Aristotle himself may have thought, this characterization is surely not a promising formulation for a modern metaphysician to start from; while both nominalists and sparse realists like Armstrong would reject the existence of Being itself outright, even promiscuous realists who might accept that such an entity exists are not likely to be favourable to Syrianus’s view that Being itself would be more worthy than anything else. Surely not only good things but also worthless and evil things exist; suffering and malevolence and other such things also exist i.e. have being, and it is surely worse that such things are than that they would not be, so being is by no means good as such. Also even Syrianus goes on to remark (see [Syr08, page 92]) that metaphysics considers (besides this Being itself) all beings, calling metaphysics a ”science of all beings”. Surely even a philosopher who does not think that this Being itself exists, or that there is much to say about it, can still consider all beings as beings.
It follows that metaphysics or ontology in the sense of the science (or
discipline) concerned with being as being may be possible even if there are
no separate substances and no gods or angels. Therefore even atheists and
agnostics can accept that ontology in this sense is possible even if they do
not accept the possibility of natural theology. Of course someone who ac-
cepts this definition can still accept natural theology and even view it as
closely connected to ontology (as Wolff himself obviously did); the point
is that this definition of ontology is neutral between religious and irreli-
gious philosophies. It is no wonder then that many, perhaps most religious
philosophers such as neo-scholastics (e. g. Etienne Gilson, Joseph Owens
and Philip Merlan in [Mer53]) typically oppose the separation of ontology
and natural theology, viewing it as a corruption of Aristotle’s original no-
tion of metaphysics (though it is rather curious that in this they oppose
the medievals, who otherwise are their models). For instance, Owens says
in [Owe63, page 25] that

in Aristotle an ontology is impossible. "Ontology" is here under-
stood in its historically authentic meaning of a general science of
Being qua Being that is in some way, at least partially, distinct
from a philosophical theology.

However, this opposition is not universal even among neo-scholastics; Józef
M. Bocheński thought contrary to them in [Boc74, page 284] that Wolff’s
conception of ontology was quite in the Aristotelian spirit.\textsuperscript{11} However, there
is no historical reason to deny that this separation is at least as natural a
development of Aristotle’s conceptions as their own. Moreover, if religious
philosophers want to have a metaphysical debate with irreligious ones, then
they must be willing to accept a neutral starting point such as the definition
of ontology as a science of being as being offers. We may even suspect that if
ontology is to prosper (instead of being confined to a small stifling ghetto of
true religious believers) or perhaps even survive in the multicultural, liberal
culture hopefully coming into being today, then such secularization may be
necessary.

\textsuperscript{11}Bocheński was also a neo-scholastic, but more logically oriented, being part of the
Cracow Circle that tried to combine Neo-Thomism with modern logic. This Cracow Circle
was in combining modern analytical philosophy with Aristotelianism an often neglected
precursor of the Neo-Aristotelian current in modern analytical metaphysics, though of
course most modern Neo-Aristotelians do not share their commitment to Catholic religion
and base their Aristotelianism directly on Aristotle rather than on Thomas Aquinas.
There may of course be significant differences between Aristotle’s understanding of the task of wisdom and Wolff’s conception of the task of ontology. To take perhaps the most important and famous example, it seems that Wolff understood ontology’s task as primarily the study of possible being (inspired by Leibniz’s use of the conception of possible worlds derived from the Scotist tradition) while Aristotle probably understood it as primarily the study of actual being (though this is also very debatable, and depends on how his obscure doctrine of potentiality is interpreted). "Being,” Wolff says in [Wol30, §134, page 115], "is what can exist and, consequently, that with which existence is not incompatible":

Ens dicitur quod existere potest, consequenter cui existentia non repugnat.

This makes Wolff’s ontology in some respects a precursor of such modern possibilist metaphysical schemes as David Lewis’s; however, Wolff of course does not propound anything like Lewis’s indexical analysis of actuality, so his possibilism may have been less extreme than Lewis’s.

Later ontologists have often gone so far as to take this concern with possibility as a defining characteristic of ontology. This includes both early successors of Wolff such as Christian August Crusius (1715-1775) in [Cru66] and also relatively recent ontologists.

2.1.3 Recent History of the Concepts of Metaphysics and Ontology; Heinrich Scholz and the Phenomenological Tradition

One of the most important though undeservedly little known metaphysicians in the early 20th century was Heinrich Scholz (1884-1956), founder of what is called the Münster Group. He belongs to those who took the concern with possibility as a defining characteristic of ontology. After distinguishing ontology as a type of metaphysics from theology and cosmology (here unconsciously following Wolff) in opposition to Kant Scholz says in [Sch41, pages 13-14] that metaphysics is ontology in the sense of the theory which contains all truths concerning entities which can meaningfully be understood as individuals which can be formulated in the language he was using that are not restricted to any domain of individuals or world, but are of unrestricted
validity, i.e. hold in every non-empty domain of individuals and in this well-determined sense hold in every possible world.\textsuperscript{12}

Scholz thought that modern logic can be taken as an ontology in this sense; he develops the theory of identity (what would be considered today as a part of predicate logic or its metalogic) as an example of a metaphysical theory. This may sound strange, as logic and ontology are usually sharply separated today; however, it has historical precedents. As Scholz himself remarked in [Sch41, page 146], Aristotle already thought that many subjects that we would today call logical (or possibly metalogical or part of the philosophy of logic) are part of the subject matter of metaphysics. For instance, he thought in \textit{Metaphysics} 1005a19-1005b18 that metaphysics studies such axioms as the Principle of Non-Contradiction (also called the Law of Contradiction) or the principle of the excluded middle and tries to justify them so far as and in the way that so fundamental principles can at all be just-

\textsuperscript{12}Scholz denies that metaphysics is the science of being as being, though his notion of unrestricted validity seems to be a possible way to interpret that very notion of concerning being as such. Scholz says in [Sch41, pages 142-145] that his conception of metaphysics is a Leibnizian one, though he admits in [Sch41, page 143] that Leibniz combined it with an entirely different, theological metaphysics. However, Scholz denies in [Sch41, pages 149-150] most similarities between his theory and Wolff’s, though for rather unconvincing reasons; he admits some similarities to Wolff’s philosophical cosmology, not to Wolff’s ontology as would be more natural.
This passage is very hard to reconcile with an interpretation that would restrict Aristotelian metaphysics to theology, for it is hard to see how the Principle of Non-Contradiction would be any more relevant to theology than to special sciences like physics. It is also hard to reconcile even with Aristotelian formulates the principle of non-contradiction. He says (in Tredennick’s translation) that

It is impossible for the same attribute at once to belong and not to belong to the same thing and in the same relation.

Today many logicians would rather formulate the principle so that it would say that the same sentence cannot be both true and false. These different formulations of the principle appear to affect to what science or discipline the principle seems to belong. One can argue that these different formulations are not really formulations of the same principle but different principles; the later principle is a semantical principle and not an explicitly ontological principle, and it is of course also possible to formulate the principle in ways that make it seem syntactic, as is commonly done in logic (though this threatens to trivialize the principle completely). Aristotle’s formulation, however, makes the principle a principle concerning entities in general, so that it does not speak explicitly of sentences or other linguistic or psychological objects at all. Aristotle’s formulation may seem to be too narrow (as it e. g. does not seem to cover the claim that the same polyadic relation cannot hold between the same entities in the same respect, which is naturally held to be a consequence of the Principle of Non-Contradiction), but the more general formulation that says that a proposition (in a non-linguistic sense as the content of sentences and judgments) cannot be both true and false can also be held to be a metaphysical and especially ontological principle. It might then be possible to try to separate metaphysics and logic by trying to claim that only the semantical principles are logical. However, this proposal faces many very serious difficulties. There is little agreement among logicians and semanticists about whether logical laws can be understood as semantical ones. Carnap famously claimed that logic was a part of semantics in [Car48, §14, page 60]. However, the founder of modern semantics, Alfred Tarski, on whose work Carnap based his own theory, said in [Tar44, page 354] that the semantical laws of contradiction and excluded middle should not be identified with the logical laws of contradiction and excluded middle. Clearly if this is correct, the same can be generalised to all logical laws. I tend to believe Tarski in this matter rather than Carnap. According to Tarski the logical laws belong to the sentential calculus, the most elementary part of logic. Tarski’s mention of the sentential calculus might be taken to suggest that the logical laws could be identified with syntactical laws and this would also make it possible to distinguish them from the metaphysical laws of contradiction and excluded middle that Aristotle talks about; however, that the logical laws would be purely syntactic does not seem plausible to me either. Furthermore, there are deeper difficulties. Though there probably are different principles of non-contradiction, they are obviously very intimately connected. The explicitly ontological formulation can be derived from the semantical formulation by just applying Tarski’s T-schema (together with some elementary applications of propositional logic and predicate logic) and from the syntactic formulation by giving all signs occurring in them their ordinary interpretations. Other logical principles such as the Law of Excluded Middle are also formulated in these different ways, and the same seems to hold of most if not all logical principles. I think that this shows a very intimate connection between the sciences or disciplines of ontology (and hence metaphysics) and logic.
an interpretation that would restrict Aristotle’s metaphysics to aetiology (as for example Wehrle’s interpretation would do) for it is hard to see how the principle of non-contradiction would explain anything in the world or give the causes of anything (even in the rather broad sense in which Aristotle uses the word αἰτία aitia, cause). Of course, the principle must hold of the primary causes and principles, but no more than of all other entities. Rather this passage in Aristotle gives support for the metametaphysical thesis of mild deflationism; however, ultimately it only gives support that the thesis applies only to some metaphysical claims (as I held in the introduction), since as we have seen, other claims made by Aristotle (such as the claim of the existence of immaterial intelligences) cannot plausibly be treated in this way. However, if the principle of non-contradiction is taken to be metaphysically significant, then there is no reason not to take syllogistics, which would today be called predicate logic, to also possess the same kind of metaphysical significance if appropriately formulated, and since the theory of identity is taken by today’s logicians to belong to predicate logic (unlike Aristotelians, who would not have included it in syllogistics) it is naturally taken to have the same kind of metaphysical significance, which of course leads straight to Scholz’s position.

Most medieval logicians differed from their master Aristotle in constantly distinguishing logic sharply from metaphysics. However, Wolff followed Aristotle (even if unintentionally) in connecting principles today’s scholars would call logical with metaphysics, taking the discussion of the principle of non-contradiction as a part of ontology, and indeed beginning the actual development of his ontological system, after the prologomena, with a lengthy

belong and fail to belong to the same sensible (physical) thing in the same respect. This interpretation would make Aristotle anticipate dialetheism. It must be admitted that this view has antecedents in Plato’s thought, as Plato apparently held that sensible reality was contradictory, and Merlan stresses the similarities between Plato and the early Platonists of the Old Academy and Aristotle rather than their differences. However, this seems to me rather implausible even as a purely historical interpretation of Aristotle’s theory, and even more implausible as a basis for modern metaphysics. Merlan’s theory is an example of a rather extreme version of a developmental hypothesis, taking Aristotle’s early philosophy to be a predecessor of Neoplatonism, such as few current historians of philosophy would accept. Merlan suggests that modern philosophers would only find his interpretation strange because they are generally nominalists or semi-nominalists; however, while I defend a rather strong form of realism in this work, and am indeed attracted to a view that Merlan would label excessive realism, I yet find very strange and unacceptable the view that the principle of non-contradiction would not hold for (statements concerning) sensible particulars.
discussion of this principle, as seen in [Wol30, §27-§55]. Wolff’s definition of logic is a different matter, as it would clearly distinguish what Wolff thinks logic to be from metaphysics. In [Wol40, §61, page 30] (see [Wol06, §61, page 40] for a German translation) Wolff defines logic as the part of philosophy that concerns the use of the cognitive faculty in the knowledge of truth and the avoidance of error. However, such a definition of logic would today be taken to apply to epistemology or heuristics rather than logic or at least pure logic (as was of course also the case with regard to most definitions of logic in later times until Frege transformed the dominant conception of logic). Many of the subjects Wolff actually discusses in his logic, such as the different operations of the mind he discusses in the very beginning of his logic proper in [Wol40, §30 - §58, pages 125-142], would also be today counted as part of epistemology or heuristics, and still others ascribed to a third discipline, such as the interpretation of sacred scriptures (discussed by him in [Wol40, §968 - §981, pages 692-706]) to hermeneutics or the convincing of others (in [Wol40, §982-§1016, pages 706-733]) to rhetoric. Of course many of them, such as the syllogisms he discusses there would also today be counted as part of pure logic, so what we would today call logic was in Wolff divided between his ontology and what he called logic and mixed in both with other kinds of material. Wolff saw (see [Wol40, §89,§90, pages 39-40] and [Wol06, §89, page 52] for a modern German translation) logic as borrowing its principles from ontology and also psychology (in hindsight unfortunately, as this helped to spread psychologism, though so far as Wolff’s logic was really epistemology basing it partly on psychology may have been quite justified, as many of today’s epistemologists still think that epistemology must be partly based on psychology). Bocheński has argued in [Boc74], conformably to Scholz’s view, that there are better reasons to link ontology to modern logic than to medieval logic, since modern logic contains laws and not only rules; more recently, Cocchiarella has defended the same view as Bocheński in [Coc01].

Scholz divided philosophy in [Sch41, page 155-158] into transcendental philosophy (using the word ”transcendental” in a rather idiosyncratic sense, which he carefully distinguishes from the Kantian sense) and real philosophy, where transcendental philosophy is concerned with truths that hold in all possible worlds while real philosophy is concerned with truths concerning the actual world (and therefore cannot be as rigorous as transcendental
philosophy). Scholz admitted that there are also types of metaphysics that are a part of real philosophy and divides them in [Sch41, page 162] into the metaphysics of nature (which oddly seems to be a part of natural science rather than philosophy) and the metaphysics of human spirit; these clearly correspond in Wolff’s division of philosophical disciplines to types of metaphysics which lie outside ontology.

I must warn that I am not saying that Scholz’s conception of metaphysics would be acceptable as such. Unlike Scholz, I do not think that all statements that hold of all possible entities are logical in the strict sense; as the phenomenologists saw (see e.g. Husserl in [Hus70, Investigation II §11, pages 455-457]), we need not only formal ontology, which plausibly does coincide with logic interpreted in one way, but also material ontologies. However, I do think that logic and ontology are more closely connected than most philosophers think today. Also, the restriction to truth concerning individuals seems arbitrary. Indeed Scholz himself admits in [Sch41, pages 135-137] that his theory of identity could be generalized to a theory of identity also concerning properties of individuals or relations between them or classes (though he suggests in the spirit of extensionalism that we could speak of classes instead of properties, here anticipating Quine’s controversial views, and suggesting an answer to the Problem of Universals that is called class nominalism or the answer that is called class realism), and presumably such a theory would still be metaphysical in his sense.

Another contentious part of Wolff’s conception of ontology is that according to him it was a wholly demonstrative science; indeed, Wolff argues in [Wol30, §2, page 2] that every science is demonstrative, so metaphysics is also demonstrative as it is a science.  Aristotle did not yet have any clear distinction between demonstrative and non-demonstrative inference, so it is contentious how far Wolff was here following Aristotle. However, Wolff’s conception of sciences as demonstrative is clearly similar to the conception of science (as designated by the Greek word ”ἐπιστήµη” i. e. ”epistême”, with which the Latin ”scientia” is often synonymous) presented in Aristotle’s Posterior Analytics and accepted by most Aristotelians as well as to

15 However, there is some textual evidence that Wolff does not think that ontology or philosophy as a whole would have to be completely a priori, as Kant held the dogmatic metaphysics he criticized to be, as I will show later when discussing Kant’s theory.

16 As we have already seen, Willem de Jong calls this conception deriving from Aristotle’s Posterior Analytics the Classical Model of Science in [dJB10], and he also counts Wolff
the rationalist idea (seen famously in Spinoza) of metaphysics as proved *ordine geometrico*, in geometrical order.\(^{17}\) However, such differences of detail, important as they are, do not undermine the near synonymy of the fundamental definitions of the disciplines in question in Aristotle and Wolff. Furthermore, all features of Wolff’s conception of ontology do not follow from the definition he gave to it, so a modern ontologist can still follow basically the Wolffian definition of ontology even if he disagrees with Wolff’s possibilism or extreme rationalism.

Later in the phenomenological tradition ontology and metaphysics were distinguished from each other in a slightly different way; instead of ontology being a part (the most fundamental part) of metaphysics it was taken to be a more fundamental discipline on which metaphysics was founded. Phenomenologicalontologists took, following Wolff and in harmony with their contemporary Scholz (whether there was any influence between Scholz and the phenomenologists in either direction or not), ontology to be concerned with truths concerning possible being which could be found out a priori, while they took metaphysics to be truths concerning actual concrete being whose discovery required some empirical evidence. This distinction was made by the founder of phenomenology, Edmund Husserl himself, but Husserl did not himself develop ontology or metaphysics very far; other phenomenologists went deeper into metaphysics, with Roman Ingarden especially being the most thorough phenomenological ontologist. It seems to me that this way of dividing ontology and metaphysics might still be useful today for those philosophers who are not extreme empiricists and allow the

---

\(^{17}\) This was the part of Wolff’s ontology that Immanuel Kant memorably rebelled against. However, Kant never questioned Wolff’s assumption that if any ontological question could be answered at all it would have to be demonstratively, by means of pure reason, but tried to show that many of the questions Wolff had tried to answer could not be answered in this way, and then drew from this on the basis of the assumption the unwarranted conclusion that they could not be answered at all.
notion of possibility to be meaningful, though it is cast in doubt by Kripke’s claim that necessary statements need not be a priori nor conversely, which seems to suggest a fourfold rather than twofold division.

According to the standard view of modern logicians (developed on the basis of such simple equivalences as I pointed to), the notion of being (in the sense of existence) can be explicated with the aid of quantifiers, which - as I will show later in this dissertation - are standardly taken to be objectual quantifiers. Therefore if the standard view of modern logic is right in this, this suggests that anything that is objectually quantified over is something that metaphysics in the sense of the Aristotelian definition as the science of being as being, and hence ontology, must take account of, so that whatever one (objectually) quantifies over he is ontologically committed to. Of course, there are logicians (such as adherents of Meinongian logic) who reject the standard view concerning the connection between existence and quantifiers, and whatever evidence there might be for their position would naturally undermine my argument for a broadly Quinean theory of ontological commitment. I cannot in this dissertation discuss such views at length, though I certainly think they should be considered seriously despite their departure from this majority view, but I will in Section 5.10 shortly show how one important ostensible counterexample to the standard view can be handled.

2.1.4 Ontological Fundamentalness

Unlike the view that metaphysics is concerned with separate substances, the proposal that metaphysics would be concerned with the study of ultimate causes and principles, i. e. be aetiology, still has many defenders today, even if it appears in a slightly generalized form. In order to see this, we must notice that the concept of cause Aristotle made use of in the definition of metaphysics as aetiology was very wide, wider than the concept of cause used today (so much so, that the use of the word ”cause” as a translation of the word Aristotle uses (aitia) is somewhat misleading, though it is hard to find a better one; ”explanatory factor” might work, but then the word ”explanatory” must be used in a wider sense than often today, e. g. in the theory of abduction). Aristotelian causes included material and formal causes, which were constituents of the entities whose causes they were. We can naturally generalize this so that the ultimate causes for which metaphysics as aetiology seeks include all ultimate constituents, not just prime
matter and forms (which many current metaphysicians would not admit to exist at all). Aristotle’s word “principle” (αρχη, arkhē) has also a very wide meaning, standing for what is first in almost any kind of order, so entities which are first in the order of constituency, i.e. fundamental constituents, can also be called principles.

It is easy to find that many modern metaphysicians accept such a characterization of their discipline. E.g. Keith Campbell says in [Cam90, page 1]:

> Metaphysics is an ambitious subject; it aspires, among other things, to give an account of the fundamental constituents of any reality and an exposition of how these constituents mesh to give us the reality in question . . . To offer a description of the basic constituents in a real situation and of the relations between them, is to furnish an ontological assay of this situation.

Many other ontologists of the past century such as Gustav Bergmann have offered similar, though often slightly narrower, descriptions of the task of ontology. Bergmann says in [Ber92, page 43]:

> Ontology accounts for everything there is in terms of simples.

As I will show later in 3.3, even philosophers who explicitly denied the possibility of metaphysics such as logical positivists also asked what is fundamental so they too engaged in metaphysics as aetiology without realizing it themselves.

A big problem with the definition is its lack of clarity; it has seldom been clear, what it means to say that something is ontologically ultimate or fundamental or simple. The notion of ontological fundamentalness has been explicated with the aid of such notions as supervenience, dependence, reduction, parthood, etc. It is by no means obvious that all of these notions of fundamentalness are equivalent.

Though it does not have as many presuppositions as the previous alternative definition this definition also contains a doubtful assumption, that there are any ultimate causes and principles. However, it has never been shown that there must be any absolutely fundamental or simple entities in any natural sense of the word “fundamental”. Indeed, some philosophers such as Jonathan Schaffer in [Sch03] and Ned Markosian in [Mar05] have
explicitly argued that there are not any fundamental entities. Schaffer argues that fundamentalism is neither logically or analytically necessary, nor does it have any empirical support; indeed, induction rather supports its opposite, even if not very strongly. Entities such as atoms, which were first thought to be simple and therefore given the name "atom" (which meant "indivisible" or "simple" in ancient Greek) have later been very frequently discovered to have constituents, so we can infer inductively that it is likely that this will continue happening.

The theory that there are absolutely fundamental entities has often (e. g. by Uriah Kriegel who is for it in [Kri04] and by Jonathan Schaffer who is against it in [Sch03]) been called (metaphysical) fundamentalism. This must of course be distinguished from other kinds of fundamentalism such as religious fundamentalism. It is also sometimes called ontological or metaphysical foundationalism; it must of course be distinguished from epistemological foundationalism. The word "fundamentalism" occurring alone is more often used for an epistemological theory, epistemological foundationalism, so the danger of confusion is great if it is used without a qualifier and "fundamentalism" is a better word for the metaphysical theory. Actually Kriegel uses the word for a still stronger theory, what might better be called monistic fundamentalism, according to which not only are all entities reducible to fundamental entities but all fundamental entities are of the same kind, of one single category (in the case of Kriegel’s theory, tropes, as in the theory of Campbell). The opposite of the weaker doctrine, which Schaffer argued

---

18This is not obvious; there are infinite regress arguments that purport to prove that denying fundamentalism leads to a vicious regress which Schaffer did not address. However, I do not think the infinite regress synechism entails is vicious. Infinite regress arguments for foundationalism in epistemology seem to be valid because the human mind is finite, and cannot go through an infinite number of premises to reach a conclusion. This has indeed sometimes been denied: there is an anti-foundationalist epistemological thesis called infinitism which denies this, and the founder of pragmatism, Peirce may have first held it in [Pei68]. However, not many of the modern opponents of foundationalism are infinitists but far more often coherentists. In any case, whatever we hold about the possibility of an infinite regress in the case of justification, the case of ontological priority is not analogous to it. Even a fundamentalist can hold the universe to be infinite (as he can hold it to be infinitely big even if arguably not infinitely divisible). Set-theory certainly implies that the set-theoretic universe is infinite, and since the set-theoretic universe is a part of the universe in the wide sense in which metaphysics studies the universe in general, it follows that if set-theory is interpreted realistically, the universe in the metaphysical sense of the word must be taken to be infinite. Whether the universe of physical particulars is infinite in any respect (e. g. infinitely big or infinitely divisible or infinitely old) seems to be an empirical question of cosmology, not an analytical question as such a regress argument would presuppose.
for, can perhaps following Peirce be called synechism, though that word is also subject to different interpretations. I do not see that there are any very good arguments for either doctrine, but there are also no strong arguments against either. However, while pluralistic fundamentalism may be a viable option, monistic fundamentalism seems to me to be a completely hopeless doctrine.

There is a still stronger sense in which the word "fundamentalism" can be used; fundamentalism can be understood so that it signifies a theory according to which not only there are absolutely ontologically fundamental entities, but properly speaking there are no other entities. This view is often formulated so that it presupposes existential pluralism or existential multivocalism; fundamental entities exist in a different and more basic sense of the word "exist" than all other entities. I will examine this kind of view in 5.5. However, this view can also be formulated in a more radical (and intuitively extremely implausible) form according to which it is just incorrect to claim that anything non-fundamental exists. Stathis Psillos uses the word in this sense in [Psi05] (for a theory that he opposes); a different and stronger sense of the word than that according to which it claims that there are fundamental entities. I will call fundamentalism in the sense in which it claims that only fundamental entities exist (in the strict sense) strong fundamentalism and the weaker kind of fundamentalism which only claims that there are fundamental entities at the basis of all other entities weak fundamentalism. Strong fundamentalism is also sometimes called sparse realism in the broadest sense of that term, in which it is applicable to all entities; the word "sparse realism" is often used so that it only refers to a thesis concerning universals. I will argue against such a thesis concerning universals in more detail later. The denial of fundamentalism or sparse realism, for which I will be tentatively arguing, can be called promiscuous realism in the broadest sense of that word.

Psillos emphasizes that fundamentalism in this strong sense does not follow from scientific realism19. Many fundamentalists in the weaker sense

19Psillos distinguishes three doctrines, realism, naturalism and essentialism, where naturalism implies strong fundamentalism, as it holds that physical entities are the only fundamental entities and the only really real ones. Psillos argues that none of these doctrines follows from the others. He shows that different philosophers support different combinations of these three doctrines and their opposites. According to Psillos while Ellis is a realist, a naturalist and an essentialist, Quine is a naturalist and realist but anti-essentialist, while Musgrave is realist, anti-naturalist and anti-essentialist. The position for
such as Campbell are also fundamentalists in this stronger sense.

Many analytic philosophers, especially Bertrand Russell and Ludwig Wittgenstein, supported an extreme pluralism called logical atomism. This theory in many of its forms was a form of strong fundamentalism. Logical atomism in its extreme form holds that all entities can be analysed into simple entities, and that only simple entities are really real, really existing, the rest being just logical constructions out of them. Thus it implies fundamentalism in the sense Schaffer uses the word, and furthermore the stronger kind of fundamentalism I have called strong fundamentalism.

Simplicity need not always be here understood in terms of the notion of constituency, as one might at first think, though it often is; Carnap at least denied explicitly this interpretation in [Car67, S 36, pages 61-63], distinguishing what he called complexes (which is to be separated from Simons’s notion) from wholes. Logical atomists often understand simplicity instead with the aid of some notion of reduction (as in Carnap), and sometimes also with the aid of the notions of supervenience (as in Armstrong) or dependence in some sense of these notions that we have seen to be very ambiguous; an entity is simple iff it is not reducible non-circularly to more fundamental entities or does not supervene on such entities or is not dependent on such entities. Logical atomism also holds that every simple entity is independent of every other entity. Relations of dependence could according to it only hold between complex entities that were in the final analysis not real anyway.

Wittgenstein seems to have accepted this kind of atomism in the sense of strong fundamentalism in his earlier philosophy (though given his obscurity one can never be certain), since he believed that all complex entities (complexes) could be resolved into simple objects (Gegenstände) [Wit21, §2.02, 2.0201] and that all things are independent in so far as they can occur in all possible situations (Sachlagen; see Tractatus §2.0122). However, it is not in any way apparent from the exposition of Wittgenstein why he should believe in these peculiar theses, and I do not have the space to review the immense literature that tries to reconstruct his thought processes.

Russell did not go quite this far since he admitted that there might be no simples and was not thus committed to fundamentalism. However, Russell which I will be arguing in this dissertation is such that I am a realist and an essentialist, but not a fundamentalist and therefore not a naturalist in Psillos’s sense.
came quite close to this extreme atomism, to strong fundamentalism. He did assume that entities such as material bodies that would ordinarily be thought to be real were logical constructions out of simpler entities, and that therefore one did not have to suppose they really existed. Stathis Psillos called this view fundamentalism in [Psi05]; this is clearly what I have called strong fundamentalism, a different sense of the word than that according to which there are simples.

However, if there are no simples as Russell allows may be the case then Russell’s assumption is cast in a weird light. Russell tries to have strong fundamentalism without weak fundamentalism and it is easy to show that this is not a coherent view. Let us suppose that instead of there being simples entities of level 0 can be constructed out of entities of level -1, entities of level -1 of entities of level -2, etc. In this case by Russell’s principles we do not have to suppose that there are any entities of level 0, since they are mere constructions out of level -1. Nor do we have to suppose there are entities of level -1, since they can be constructed out of entities of level -2, etc. In short, we do not have to suppose there is anything at all! This is surely absurd. Something has clearly gone wrong, and I suggest that what has gone wrong is the assumption that we do not need to suppose that complex entities really exist.

Strangely enough Schaffer in a later article [Sch09] has changed his opinion (without explicitly acknowledging that he held a different view) and how holds that there is a fundamental entity; in fact he tells in [Sch09, page 361] that there is exactly one fundamental entity, the whole concrete cosmos.

Bergmann says in [Ber57, page 325] that philosophers who maintained that bodies did not really exist (as Russell surely did) were either raving mad or were using the word “exist” in a peculiar philosophical way. While it is surely legitimate to use words in unusual technical senses in philosophy. I do not see any utility in using the word “exist” is such a peculiar sense. If in saying that only simples really exist we use ”exist” so that it means ”exist and are simple”, the thesis becomes the truly trivial thesis that only simple entities which exist exist and are simple. Nor do I see any real reason to deny that Russell meant what he said.

Schaffer himself acknowledges that this is form of monism; however, he distinguishes it as priority monism from the more common kind of existence monism, and argues that it is more plausible than existence monism. I admit that priority monism is more plausible than existence monism, but do not find either kind of monism very plausible. While I think that there probably are not any fundamental entities, yet if there were, I think that would be far more plausible to hold that there are many fundamental entities than one single fundamental entity. Schaffer’s arguments for priority monism seem rather weak to me; he argues in [Sch09, page 378] for monism on the basis of the assumption that the grounding relations are relations of abstraction. His argument for this is that the derivative entities,
However, it is not clear to me why Schaffer changed his mind, and as far as I can see he gives no good argument for his new view that there is a fundamental entity, or even any counter-arguments to his earlier arguments that there are no fundamental entities. Schaffer does refer to Aristotle’s example, but just appealing to Aristotle’s authority is not a good argument. I am more persuaded by his earlier argument against the view that so far as we can know there would be fundamental entities. Schaffer holds in [Sch09, page 376] that the relationship of ontological priority is unanalyzable; however, this view is vulnerable to the suggestion of the opponent of metaphysics that the notion would be meaningless. Schaffer argues in [Sch08a] that priority seems no worse than notions such as parthood and causation; however, there is a difference. The notions of parthood and causation are frequently used outside ontology; however, it is not clear whether the notion of ontological priority is ever used outside of ontology. Because of this it can be argued that the burden of proof is on one that claims that the notions of parthood or causation are meaningless; the claim is implausible on the face of it. However, it is more plausible to claim, as a positivist or other enemy of metaphysics is inclined to, that the notion of ontological priority is meaningless.

Therefore the tasks of metaphysics or ontology described by Campbell and Bergmann may not be attainable ones, not only because of any limits to human epistemic capacities but because they are intrinsically impossible. If all entities do not consist of absolutely fundamental entities you cannot reduce them to such entities no matter how clever you are. Thus it may be intrinsically impossible to furnish an ontological assay in Campbell’s and Bergmann’s sense. Of course even if there were fundamental entities and so the task was intrinsically possible, it might still be impossible for human beings.

This conception of the task of metaphysics is apparently not equivalent in order to be an ontological free lunch and count as no further addition ought to be already latent within the substances (by which word Schaffer just means the fundamental entities). I find plausible the premise that the derivative entities ought to be a free lunch (though I do not find it at all certain). However, it seems to me that (as many philosophers such as Armstrong and Lewis have already argued) the derivative entities count as no further addition also if the grounding relations are relations of composition; however, this assumption rather suggests a plurality of fundamental entities. On these grounds at least the monistic and pluralistic kinds of fundamentalism are equally plausible (though as I have argued, I do not think there are very good arguments for either of them).
with the proposal that ontology is the study of being as being, at least if the idea that being is expressed by the particular i.e. existential quantifier is correct, since many beings that we quantify over existentially are surely not ultimate causes or principles in any obvious sense. However, these proposals might be equivalent if we postulate that we only ought to quantify over fundamental entities. However, it is controversial, to say the least, whether this postulation (which has been called fundamentalism in a strong sense of that word) is correct or useful in any way. Schaffer quite plausibly rejects this view in both of the stages of his philosophical development.

However, it does not follow that there is nothing to the way Campbell and Bergmann formulate the task of ontology. Even if there are no fundamental entities, we can yet ask, as Schaffer in [Sch09, page 376] does, what grounds what. It is quite plausible that this kind of question is one of the basic kinds of questions in metaphysics or ontology; Schaffer gives rather persuasive arguments for the view that this kind of question belongs to metaphysics. It does seem intuitive that ontology would be more interested in fundamental entities than superficial ones. Ontology does somehow try to get below the surface of the world. Unlike what Quine’s definition might suggest at first glance, ontological research does not of course consist simply in listing entities or classes of entities at random like: there are cabbages, kings, turnips, ashtrays, suns, John Kerry, New York, blades of grass, Quine, etc. Ontology tries to give a systematic answer to the question of what there is. One way of doing this is of course classification; ontology tries to find out general classes or kinds of things such that all entities can be assigned to them. The most general of these kinds or classes are often called categories. However, classification need not suffice to make ontological categories. Just how categories differ from other classes or kinds is very controversial. Classically (e.g. by Aristotle) it has been thought that categories are the absolutely most general kinds so that no category could be a subclass or subkind of any other class or kind. However, it seems to me that this would make the notion of a category so narrow as to be virtually useless. Even Aristotle, though he apparently thought there were only ten categories, none of which overlapped another, made use in his theory of categories of the notion of accidents. This raises the question why could not accident be a genus and a category of which nine of the categories would be subclasses and hence subcategories? However, surely metaphysics must distinguish between these subcategories if it is not to be utterly trivial. Also in his metaphysics (1029a) Aristotle said that in one sense the substance is matter, in another the form and in a third a combination of the two, and it would seem that in a total Aristotelian system the categories of matter and substantial form and the combination of both should be subcategories of the category of substance. One answer, which is also based on the thought of Aristotle, is that categories represent...
enquiry sufficiently systematic. Two kinds of entity might be interestingly related even though they are mutually exclusive; a man and his heart are thus related, though no heart can be a man nor any man a heart. A man has a heart as a part and is dependent on a heart so that he cannot survive without it. Any systematic theory of the world should take account of such relations as parthood, dependence, supervenience, etc. Such relations, however, can be used to divide the world into less fundamental and more fundamental entities. Therefore telling what fundamental entities there are and how the rest can be derived from them does at first sight appear to be an attractive way of performing the task of systematically describing the world. However, it may not in the end be a feasible way of performing that task. Yet perhaps it can be modified so that it becomes more certainly feasible.

We can formulate more modest versions of the ontological task. Even if there are no absolutely fundamental or simple entities, there can be entities that are more fundamental or simpler than any other entities that are known or knowable to us. As was already pointed out by Eugene Bronstein in [Bro34], even if some things are simpler than others and can be used to analyse the less simple things it does not follow that anything would be absolutely simple. Even this assumption is unnecessarily strong, however. Even if there are no such entities, there may yet be entities that are at least as fundamental as any other entities known to us; i.e. all other entities are reducible to them, though they themselves may be reducible to others. Even when two kinds of entities are equally fundamental, we can arbitrarily take one of them as basic in constructing an ontological theory. We can say that the aim of ontology is to reduce all entities known to us to some set (which should be as small as possible) of entities at least as fundamental as any other entities known to us, even though the entities in such a set might themselves be derivable from yet further entities not known to us and the set might not be unique. Such a reduction helps us to answer the question of what there is more systematically than a mere classification by means of different modes of existence. The use of categories in this sense would presuppose a theory called existential pluralism or existential multivocalism, which I will later argue against in Section 5.5. The way to the correct answer seems to start from the idea that that categories are more general than any class whose concept is specific to a science; categories are kinds either used by all sciences or too general to be used by any discipline other than metaphysics. This seems to allow a category to be a subcategory of another. However, it is not yet clear whether such an explanation can establish any sharp boundary between categories and other classes.
categories. An account of such entities that are at least as fundamental as any others may therefore be called an ontological assay in a weakened sense.

Schaffer argues persuasively in [Sch09, page 354] that the target of metaphysical inquiry should be an ordered rather than a flat structure. Instead of just finding a set or class or plurality or list of entities $E$, which Schaffer calls a flat structure, we should in his view seek for a pair $⟨F,G⟩$, consisting of a set $F$ of fundamental entities and a set or class $G$ of grounding relations. Schaffer argue quite plausibly that flat structure is strictly weaker than ordered structure. If we reject the view that there are fundamental entities, we cannot use this kind of ordered structure. However, this does not mean that we would have to use a flat structure. We could make use of another kind of ordered structure. We could use a set $⟨E,G⟩$ consisting of a set $E$ of entities and a set $G$ of grounding relations, even if $E$ has no minimal members, as we can have the set of negative and positive integers ordered by the lesser than relation even though there is no first negative integer, as a pair $⟨Z,<⟩$.

However, there is another reason for doubting that this will work, but this has nothing to do with whether there are fundamental entities. There are very strong arguments for the conclusion that there are too many entities for them to form a set. In most set theories, including the standard Zermelo-Fraenkel (ZF) theory, there cannot be a universal set, so if such set theories are correct $E$ cannot exist as a set in the first place but must be a proper class or not a class at all. In this case grounding relations cannot form a set either and since a proper class cannot be a member of an ordered pair in standard set theory, this gives us a reason to think that neither an ordered structure such as $⟨F,G⟩$ nor an ordered structure such as $⟨E,G⟩$ could exist, even if there were fundamental entities. Indeed, according to

---

There is one weakness in Schaffer’s argumentation. He claims in [Sch09, page 355] that given a list of entities, there is no guarantee that one can sort or order them. However, in set theory there is an axiom called the Axiom of Choice, which asserts that any set of entities can be ordered. Most mathematicians and logicians find this axiom intuitively plausible, which gives us a reason to think that the set of entities belonging to any list can be ordered. Indeed, the use of the word “list” suggests that the entities are already ordered in some way, as the word is commonly (especially in computer science) used for ordered structures. However, all orderings are not equally good for cognitive purposes.

In fact it is dubious whether the target of metaphysical inquiry is usefully conceived as a set of entities at all. I would say that what metaphysics, or any scholarly discipline strives to arrive at, is a set of propositions, or, since propositions have to be expressed in sentences, a set of sentences, not a set of entities, ordered or unordered.
set theories which not only do not allow a universal set but do not make any
distinction between sets and classes either even a set such as $E$ itself could
not exist, so there are reasons to doubt whether $E$ exists at all. However,
ordered structures such as $\langle E_i, G_i \rangle$ could exist where $E_i$ are the sets of
entities belonging to some special sufficiently small category or class, and
$G_i$ are the grounding relations restricted to this set; for example, Schaffer’s
concrete objects might be such a special category, and even if they are not
some subclass of that category surely would be. However, there probably
are too many of them so that there cannot be any set which would contain
all of them. Of course, if there were fundamental entities, then a structure
of the form $\langle F_i, G_i \rangle$ might also exist, where $F_i$ is the set of fundamental
entities belonging to the sufficiently small subcategory $E_i$; however, since
no good reason has been given why there would be fundamental entities, it
is better to be more cautious and use structures of the form $\langle E_i, G_i \rangle$.

2.2 Kant’s Metametaphysical Attacks on (Trans-
cendent) Metaphysics

Unfortunately, metaontology has often been used and still is often used
purely negatively, to attack ontology (as also metaethics has been used to
attack normative ethics). I must stress that I do not want to use it in such
a way, but rather positively, to find methods that can lead to more rigorous
research in ontology. Most of this dissertation will be dealing with this pos-
itive, constructive task. However, a metaphysician or an ontologist is not
justified in evading arguments that have been given against the possibility of
metaphysics, but must show that such arguments are unsuccessful. There-
fore I will in this and the following section deal briefly with metaontological
questions regarding the very legitimacy of ontology or metaphysics. These
questions turn out to be entwined with questions of the correct method of
metaphysics, for many methogological principles, most prominently both
verificationism and structuralism, have been used by different philosophers
as on the one hand tools against the very possibility of metaphysics and
on the other hand as pointers to true methods or even true contents of
metaphysics.

I must stress that the adoption of a critical method does not imply the
adoption of any of Kant’s very sceptical conclusions, such as the claim that
we cannot have any knowledge (or even justified beliefs, since in Kant’s view justified belief had to be based on knowledge) about things in themselves. In fact Kant’s own reasons for his sceptical conclusions are no longer as such relevant in today’s philosophy, and have not been so for more than a century, though there are lots of popularizers of philosophy who are out of date and still assume that Kant has shown metaphysics in the traditional sense to be impossible. Indeed, Kant is sometimes even thought to have shown all metaphysics to be impossible.

In fact Kant did not even attempt to do that; while he argued that metaphysics in the traditional sense, what is sometimes called transcendent or speculative\(^{25}\) was impossible, he thought that metaphysics in another sense, what is sometimes called immanent metaphysics, as a general doctrine concerning all possible (objects of) experience, was possible and that he was contributing to it\(^{26}\); it is only what is often called transcendental metaphysics, general doctrine concerning things in themselves, that he objected to.

Kant based his metaontological conclusions upon such epistemological claims as that metaphysics along with mathematics depends upon synthetic a priori knowledge and we can have a priori synthetic knowledge only about phenomena, since a priori synthetic knowledge is based on the forms of perception and understanding and hence could only concern phenomena\(^{27}\).

\(^{25}\)Anjan Chakravartty uses the phrase ”speculative metaphysics in [Cha07, page 16] for the kind of metaphysics that concerns unobservables. He distinguishes this speculative metaphysics from metaphysics understood most broadly, as the study of being qua being (for which he does not give any special name). I will discuss this distinction at length later in Section 3.3. Chakravartty includes speculation about universals in speculative metaphysics, but I will argue later that some universals (in the widest sense of the word) are observable, so (the most basic part of) the theory of universals does not belong to speculative or transcendent metaphysics.

\(^{26}\)Sami Pihlström argues in [Pih09, page 3] on behalf of a modern version of such a Kantian conception of metaphysics, metaphysics as a study of the basic features of a humanly categorized reality. Pihlström says explicitly that more than anyone else it was Kant who paved the way for this kind of rethinking, though he is also strongly influenced by American pragmatists (which is why he calls his theory pragmatist metaphysics) and by Hilary Putnam among modern philosophers. I do not think that this is a good conception of metaphysics (as I will partly argue later); nevertheless, it is important to notice that it is a genuine conception of metaphysics, which supports the claim that metaphysics need not be rejected even from a broadly Kantian perspective.

\(^{27}\)Exactly what these forms of perception and understanding are and how Kant supposes synthetic a priori truths to be based on them is perhaps the most obscure point in Kant’s philosophy and I do not want to delve into it at any great length. Fortunately it is not necessary to consider them since as I will show there are many dubious points in those parts of Kant’s argument against the possibility of knowledge concerning things in themselves.
From these claims it followed that all our knowledge, metaphysics included (so far as it is at all possible), can only concern phenomena. However, these Kantian epistemological doctrines are no longer viewed as plausible by many philosophers. Philosophers today either think that synthetic a priori knowledge is not possible (and that metaphysics can be based either on analytic a priori knowledge or synthetic a posteriori knowledge or a combination of both) or that it is possible but what makes it possible is something entirely different than what Kant thought it was (which opens up the possibility that such knowledge might concern things in themselves after all).

Matters are complicated by the fact that the notion of synthetic statement and the notion of experience involved in the definition of a priori knowledge can be (and have often been) argued to be unclear and ambiguous and Kant’s claims might be true in one sense but false in another, and therefore their implications are not as clear as might have been thought. While Kant’s arguments were epistemological he also made use of notions that would later have been classified as semantical, such as the notions of analytic and synthetic, and it has been argued that he did not distinguish these notions clearly enough. Kant’s arguments depend on a specific view of the interrelationships of these two kinds of notions, which has come under attack in later philosophy. Synthetic knowledge a priori might be possible in one sense but impossible in another, and if different statements that have been proposed as ostensible examples of synthetic statements known to be true a priori are synthetic and non-empirical in different senses, then the reason for the possibility of knowing their truth might also be different. A lot of discussion in analytical philosophy has concerned just these notions, and it is obviously not possible to go at all deeply into these questions in this work.

A statement may be said to be analytically true in the widest sense if and only if it is true in virtue of the meanings of the terms it contains, but it is far from sure that this is unambiguous (”meaning” being a notoriously ambiguous term) and far from clear just what it involves. Kant tried to define the notion of analyticity more rigorously by saying (e. g. in *Critique of Pure Reason* A6-7, B10-11) that a statement is analytical if and only if its predicate concept is contained within its subject concept (and synthetic otherwise). However, not only is it unclear what the concept of containment that are slightly clearer.
used here exactly means and therefore what it involves - and it has in fact
been argued (e. g. by Jerrold Katz) that there are different senses in which a
concept could be contained in another concept (and indeed different senses
of the very word ”concept”) - but modern logic has often been taken to
imply that all statements are not of a subject-predicate form. If this is
correct then this definition does not enable us to classify all statements
as analytic or synthetic as Kant thought. Frege said in [Fre84, §3, page 4]
(see [Fre50, §3, page 4] for an English translation) that if in finding the proof
of a proposition we come only on general logical laws and definitions, then
the truth is an analytic one. Frege said in [Fre84, §3, page 3] (see [Fre50,
§3, page 3] for an English translation) that he did not mean to assign a new
sense to these terms, but only to state accurately what earlier writers, Kant
in particular, had meant by them. This is generally taken as the de facto
definition of analyticity in modern analytic philosophy, though there have
been philosophers, such as Jerrold Katz, who have said (e. g. in [Kat72,
page 119]) that Frege’s definition did not catch the way in which the term
had been used earlier, as Frege intended. It seems to me that since Kant’s
formulations are likely to be ambiguous, there may be no unique way to state
what Kant had meant by them, but Frege’s definition captures at least one
legitimate interpretation of them.

2.2.1 Four major Epistemological Positions and their metaontological Consequences

Following Quine in [Qui53e] many philosophers have reacted to this unclarity
of the notion of analyticity by going so far as to reject the very notion of
analyticity. However, if the lack of a fully satisfactory definition of a concept
were a sufficient reason to reject that concept we would be left with no
concepts to use. Quine tries eventually to explicate the notion of analyticity
by appealing to the notion of verification; however, as I will argue in the next
section, there is no reason to suppose that the notion of verification (and
the notion of experience in terms of which it is commonly defined) would be
any clearer than the notion of analyticity or of synonymy that Quine finds
problematic.

Many naturalistic philosophers (e. g. Devitt in [Dev13] or Putnam
in [Put68]) hold the radical empiricist view that no a priori knowledge (even
analytical) is possible but instead all knowledge, even knowledge of logical
truths, is a posteriori - a view which had already been held by John Stuart Mill in a cruder form in [Mil72]. Obviously if this is correct then Kant’s argument against the possibility of knowledge of things in themselves fails. Philosophical views containing this thesis are often called naturalistic, in one of the senses of that popular and annoyingly ambiguous term. While many such naturalists are of course opposed to metaphysics, this does not follow from naturalism as such and not all naturalists hold this view. Many naturalists are naturalistic metaphysicians28 who think that metaphysical truths are synthetic a posteriori like the natural sciences with which they think metaphysics is continuous (unless they follow Quine in [Qui53e] in denying the distinction between analytic and synthetic statements entirely, in which case they merely think that metaphysical truths are a posteriori like the truths of natural sciences). Some might even hold that they are analytical truths a posteriori, something that both Kant and Frege agreed, despite their enormous differences, it was impossible any truth could be.

The phrase ”naturalistic metaphysics” is seductively ambiguous, as Susan Haack has shown is the parallel phrase ”naturalistic epistemology”. Not only can ”naturalism” and ”naturalistic” refer to either a metaphysical or an epistemological view or characterization, but they remain ambiguous even if taken to refer only to an epistemological view. Susan Haack has disambiguated the phrase ”naturalistic epistemology” in chapter 6 of [Haa09], and distinctions similar to the ones she makes can also be made with regard to research programs and methodologies called naturalistic metaphysics. Haack distinguishes from each other on the one hand reformist and revolutionary naturalism and on the other hand aposteriorist and scientistic naturalism and also broad and narrow varieties of all of these kinds of naturalism. Haack’s definitions can be applied to naturalistic metaphysics if her reference to cognitive science, which is obviously the branch of natural science (if indeed it is a natural science) most relevant to epistemology, is changed to a more broad reference to natural sciences more generally. While some philosophers, such as Alvin Goldman in [Gol07], have suggested that cognitive science is relevant not only to epistemology but also to metaphysics, it is more common to claim that physics is relevant to at least the most general

28Sometimes naturalistic metaphysics is identified with analytical metaphysics or seen as a subspecies of it, but some metaphysicians, at least quite clearly James Ladyman, Don Ross and their colleagues in [LwDSC07], hold the two kinds of metaphysics to be incompatible.
and basic metaphysical problems. Aposteriorist metaphysical naturalism would claim that traditional metaphysical problems can be resolved a posteriori, with the aid of experience, while scientistic metaphysical naturalism would claim that metaphysics is internal to the sciences. If this aposteriorism is narrow, this applies only to some traditional metaphysical problems, while if it is broad it applies to all traditional metaphysical problems. Reformist naturalistic metaphysics would claim that the results of natural sciences, especially physics, are relevant to traditional metaphysical problems while revolutionary naturalistic metaphysics would claim that natural science can by itself resolve traditional metaphysical problems. The word "scientism" is generally used as a derogatory term, but at least James Ladyman has proudly proclaimed himself (along with Don Ross and David Spurrett) a defender of scientism in the first chapter of [LwDSC07]. Ladyman and his co-writers are as clear examples of what we would modifying Haack’s definition call revolutionary ontological naturalism as could be imagined. They argue in [LwDSC07, page 1] that metaphysics should be motivated exclusively by attempts to unify hypotheses and theories which are taken seriously by contemporary science. It is not wholly clear, however, what view they would take of mathematics, e. g. whether they would accept Quine’s view of it.

I will argue in favour of a modest naturalism in metaphysics, similar to Haack’s position in epistemology. I will argue that according to the traditional conception of metaphysics, some metaphysical statements are analytic a priori while others are synthetic a posteriori. This implies that at least if this traditional conception is coherent, and if Quine is not right in dismissing

\[29\] It is not clear if they would also reject pure mathematics, to which analytical metaphysics is quite similar. Would they claim that mathematics should be motivated exclusively by its applications in physics and other natural sciences? As we will see later, Ladyman and Ross are structuralists. However, since (at least most) structuralists reduce physical objects to mathematical, specifically set-theoretical structures, they must admit mathematics, specifically set theory, to be very important for science. I will later show that there are many problems in the epistemology of mathematics, and especially the epistemology of set theory, which make it still quite doubtful whether empiricism can account for mathematical knowledge. However, if it cannot then it cannot account for all of scientific knowledge either, as science requires mathematics (especially on a structuralist account of it). If a priori reasoning is then valid in mathematics why might it not be in analytic metaphysics (especially if, as I argue in this work, analytic metaphysics actually overlaps with logic and mathematics)? This old argument against scientism is far from conclusive, as it is yet far from certain that a holistic kind of empiricism cannot account for mathematics, but it at least suffices to show that scientism cannot yet be held to be conclusively established either
entirely the distinction between analytic and synthetic statements, then narrow aposteriorist metaphysical naturalism is correct but broad aposteriorist metaphysical naturalism is not. Hence scientific metaphysical naturalism is mistaken, though reformist metaphysical naturalism might be correct.

Kant’s own reasons for thinking that metaphysics would have to be synthetic a priori (as a whole) are not very impressive. In *Prolegomena* §1 A 23-24 (translated in [Kan85, §1, page 111]) Kant argued plausibly enough that the peculiar features of a science may consist of a simple difference of object, or of the sources of cognition (Erkenntnisquelle), or of the kind of cognition, or perhaps of all three conjointly. Kant went on to claim that as regards the sources of metaphysical cognition, its very concept implies that they cannot be empirical. He clarified this claim, telling that it implies the following:

- its principles (including not only its basic propositions but also its basic concepts) must never be derived from experience.

This is a very strong claim, for it consists of two logically independent claims (although this is not always seen). There are four logically consistent views about the semantic and epistemological status of a discipline; it could be held that both its basic concepts and propositions must be derived from experience, it could be held that its basic concepts must be derived from experience but its basic propositions need not be; it could be held that only its basic propositions, not its basic concepts, would have to be derived from experience and finally it could be held that neither have to be derived from experience. These four views then range from the strongest version of empiricism concerning the objects of the discipline in question to the strongest rationalism concerning these objects. Kant seems to have held that only the first and last of these four positions are coherent, though it is not clear why he thought so; at face value all four positions are coherent (and may indeed have been held by famous philosophers; as I show soon below there are good reasons to suspect that Aristotle or at least some Aristotelians implicitly held the second position with regard to some disciplines). Most later empiricists have also thought so; logical positivists in their verificationist phase or mood (as distinct from their coherentist one) commonly held that the first view held of all disciplines consisting of synthetic truths.

---

30I mean by this that knowing them would not require any additional experience beyond that involved in acquiring the concepts.
Neither is it clear why Kant held the fourth view about metaphysics. It seems to be connected with the way Kant distinguished metaphysics from physics; Kant said immediately afterwards:

It must not be physical but metaphysical knowledge, i.e. knowledge lying behind experience.

It is far from clear how Kant gets from the proposition that metaphysics is not physical knowledge to the claim that it is knowledge lying behind experience\textsuperscript{31}. If Kant’s claim that the very concept of metaphysics implies that the sources of metaphysical cognition cannot empirical is to be usable in the critique of metaphysics, then the concept of metaphysics he uses must be one which proponents of metaphysical theories themselves used, or else Kant commits a fallacy of equivocation and attacks a straw man. However, proponents of metaphysical theories have naturally had many similar but not identical concepts of metaphysics, so comparing Kant’s concept to theirs is not as simple as it seems. If Kant’s critique were to be an adequate critique of traditional metaphysics, as he claims, it would have to apply to all concepts of metaphysics held by traditional metaphysicians. Finding out whether this is the case would involve extensive historical research. However, one good way to preliminarily test whether Kant was right in his claim about what the very concept of metaphysics implies is to consider how the concept of metaphysics was defined when it was first introduced.

Unfortunately for Kant, such considerations make it highly likely that Kant did in fact commit a fallacy of equivocation, that the concept of metaphysics he used was not a traditional one. The concept of metaphysics was first introduced by Aristotle, though it was his editors, not he himself, who gave it the name ”metaphysics”, while Aristotle himself used such names as ”wisdom” and ”first philosophy”. This introduction can be held to have fixed the reference of the term ”metaphysics”. While some pre-Aristotelian theories, such as Plato’s or Parmenides’s theories, can retrospectively be counted

\textsuperscript{31}Kant may have other reasons for claiming that metaphysical claims would have to be a priori. Kant thought that such claims would have to be absolutely general and that this implied that they had to be necessarily true. This association of metaphysics with necessity is indeed in accordance with older metaphysicians, as we have seen in the case of Wolff and Crusius. Since Kant thought that only a priori claims could be necessary, this may have led him to think that metaphysical claims would have to be known a priori. However, since Kripke has famously argued in [Kri72] that there are necessary truths known a posteriori, this reason is not very convincing for a modern analytical metaphysician.
as metaphysical, yet Aristotle’s theory is a paradigmatic metaphysical theory because it was used to fix the reference of the word ”metaphysics”. Therefore any account of metaphysics that would not count Aristotle’s metaphysics as metaphysics would be absurd (though of course equally an account that would only count Aristotle’s theory as metaphysics, as some of the more extreme Neo-Aristotelian accounts would do, would also be absurd). Therefore we must look at how Aristotle introduces the concept of wisdom or first philosophy.

However, when we do so, we find that the original concept of metaphysics was not such as Kant claims. As we have seen in some detail in Section 2.1, Aristotle, who first explicitly introduced the science of metaphysics, defined it with the aid of its object. He was indeed not very clear in this, since as we have seen he gave several different definitions or characterizations of what the object of metaphysics was, which are not obviously equivalent, both that it was the science of being as being (Metaphysics 1003a) and that it was the science of primary causes and principles (Metaphysics 932a) and that it was the science of immaterial entities. This allows a great degree of latitude in the definition of metaphysics, as a modern metaphysician can choose any of these three conceptions. However all of Aristotle’s definitions or characterizations of metaphysics concerned the object of the science, not the source or kind of metaphysical cognition. Aristotle’s definitions of physics also were based on the object of it, so the difference between physics and metaphysics was according to him a difference in their objects, whether it was that physics dealt with material and metaphysics immaterial objects, or that physics concerned only one domain of being (namely material objects) while metaphysics was concerned with all domains of being as such. Since Aristotle thought that metaphysics was distinguished from other sciences by its object, this would even allow an Aristotelian to coherently claim that different metaphysical claims derive from different sources of cognition and involve different kinds of cognition. As I will show, there are reasons to think that he actually at least implicitly did make this claim, and held some

\[32\] L. A. Paul also argues in [Paul12] similarly to me that metaphysics has a distinctive subject matter, not a distinctive methodology, though Paul does not reach this conclusion by a historical analysis such as I use to arrive at it. However, Paul’s view is in other respects different from mine. Paul holds in [Paul12, §2] that both scientific and metaphysical theorizing can be understood in terms of constructing of models of the world. I will instead argue that both scientific and metaphysical theorizing must make use of fully interpreted languages, so both need something more than mere construction of models.
metaphysical claims to be analytical a priori while others to be synthetic a posteriori. This would make Aristotle’s metaphysics an anticipation of the theory Haack called narrow aposteriorist naturalistic metaphysics.

Aristotle could not have consistently held that metaphysical claims would be a priori in the very strong sense of the expression in which Kant thought them to be a priori. Indeed, Aristotle himself and faithful Aristotelians had to think, contrary to Kant, that the basic concepts of metaphysics would have to be derived from experience, since Aristotle famously thought (here setting himself in opposition to the view of his mentor Plato) that all concepts possessed by a human being had to be abstracted from experience33 (compare for example De Anima 428b10, where Aristotle says that imagination is impossible without sensation and 431b where he says that the faculty of thinking thinks the forms in the images; combining these passages we see clearly that according to Aristotle we can only think of forms - which are what corresponds to concepts in Aristotle - if we have corresponding sensations). It is indeed unclear whether Aristotle thought that the basic propositions of metaphysics were derived from experience or not, as Aristotle himself never made this distinction explicitly, so his answer to this question (if he had one at all) must be inferred from other claims he makes. Unfortunately, any such inference is difficult due to the obscurity of many features in Aristotle’s theory, so that Aristotle’s philosophy allows for both more empiricist and more rationalistic interpretations34. However, whatever Aristotle’s own view (if he even had one) on this question might have been, it was not part of the very concept of metaphysics that he formulated. In fact, as I have already explained, Aristotle’s metaphysical system included claims based upon astronomy. Kant himself would have classed such claims as synthetic a posteriori (and it seem unlikely on the face of it that Aristotle would have held astronomy to be a priori, if he had formulated the question). Therefore metaphysics in the Aristotelian sense includes a posteriori

---

33There are indeed many obscurities in Aristotle’s doctrine of abstraction. These are connected with the obscurities in his doctrine of forms, on which his doctrine of abstraction is based. As I will show later, it is not clear for example whether forms were supposed to be particulars or universals. Nevertheless, it is rather clear that Aristotle’s doctrine was a kind of concept-empiricism, and so incompatible with Kant’s radical apriorism.

34Platonists notoriously had very rationalistic epistemological views, and most Aristotelians since late antiquity - medieval scholastics definitely included - were influenced by the Neo-Platonists’ combination of Aristotelian and Platonic ideas. Therefore it is quite likely that Aristotle and the earliest peripatetics were willing to let experience pay a far larger role in metaphysics than most of the later Aristotelians.
claims, whether it also includes synthetic claims a priori or not.

Kant’s claim that the very concept of metaphysics implies that metaphysical knowledge has to be a priori would then appear completely arbitrary and unjustified to someone who only looked at Aristotle’s pioneering definition of metaphysics. In fact it is of course not wholly arbitrary, as we see when we look at the historical context in greater detail, but at least to some extent based on the conception of metaphysics held by the rationalistic Leibnizian philosophers like Christian Wolff against whom Kant argued. These Leibnizian philosophers generally held that a huge amount of metaphysics was indeed a priori but also that it was analytic. They tended to count as analytic many statements with regard to which today’s philosophers would find it absurd to even consider that they might be analytic. So for all I have said Kant’s arguments might yet have been effective against such rationalists. However, the theory of a modern Neo-Aristotelian who goes back to an Aristotelian conception of metaphysics or any modern metaphysician whose concept of metaphysics is more similar to the Aristotelian than the rationalistic one (including almost all naturalistic metaphysicians) is quite unscathed by Kant’s arguments. Therefore even if extreme empiricism were true, this would not be a sufficient reason to reject metaphysics.

In fact there are reasons to think that though Kant’s conception of metaphysics may have been closer to that of Leibnizian rationalists, even they did not generally accept quite so rationalistic a conception of metaphysics as the one Kant argues against. Wolff allowed in [Wol30, §4, page 2] (see [Wol05, pages 20-23] for a modern German translation) that ontological statements can depend on undoubted (or indubitable) experience and in [Wol40, §34, page 15] (see [Wol06, §34, page 21] for a modern German translation) that in philosophy principles must be derived from experience. Wolff used the phrase "experientia indubitata", translated by Dirk Effertz as "unbezweifelte Erfahrung" in [Wol05, pages 20-23], where the very word "Erfahrung" is the same as Kant uses. If ontological statements can according to Wolff depend on experience they are then not a priori according to Kant’s definition of a priority. In any case, Wolff said in this place also that the demonstrative method is to used also in physics, so the distinction between physics and metaphysics cannot according to him depend on the demonstrative nature of metaphysics. It is of course true that it is obvious today that there are many defects in Wolff’s conception of the methodology of metaphysics (as well as in his conception of the methodology of physics); Wolff did not see that it would be necessary to use inductive or abductive reasoning in metaphysics (or that the experience on which metaphysics depends cannot always be undoubted), which can only provide probably true results, as most modern metaphysicians would think (e. g. Wolff says in [Wol06, §33, page 21] that philosophy must strive for full certainty). Wolff was definitely a strong rather than a weak foundationalist. Nevertheless, there is no difference in this point between Wolff and Kant, for Kant was as far from fallibilism as Wolff. It is then likely that the kind of metaphysicians Kant argues against are just pure straw men, though since I cannot in this work examine the complete history of rationalistic philosophy in Germany, I cannot prove it in this work; Kant’s conception of metaphysics might be found in some obscure minor Wolffian.
as impossible.

However, extreme empiricism, though fashionable, is far from universally accepted. Perhaps most analytical philosophers think (in accordance with most classical analytical philosophers, as indeed they have to think so that they can appropriately call themselves “analytical” philosophers) that analytical a priori knowledge is possible, but that synthetic is not. This would of course be of no use if if Kant’s claim (e. g. in Prolegomena §4, A36, translated in [Kan85, page 16]) that metaphysical judgments properly (eigentlich) so-called are all synthetic were correct; however, if we again look at how the concept of metaphysics was introduced, Kant’s claim appears quite implausible. Aristotle himself had (in Metaphysics 1005a19-1005b18) counted the principle of non-contradiction as a properly metaphysical principle, and this principle is according to Kant (Prolegomena §2, A 25, translated in [Kan85, page 12]) analytic, indeed the common principle of all analytic judgments. Also Christian Wolff, whose philosophy was Kant’s primary target, held the principle of contradiction to be an ontological and hence a metaphysical principle. One of Kant’s examples of synthetic judgments a priori (which were meant to be relatively uncontroversial) were arithmetical judgments; however, most analytical philosophers have thought following Frege’s logicism that such judgments are reducible to logical statements and are hence analytical a priori, and this is probably still the standard view, though

36There were some exceptions even among earliest analytical philosophers. One exception was Russell. He indeed agreed with Frege that arithmetical statements are logical and did much to make this the common view. However, unlike Frege Russell held at least initially in [Rus03, §434, page 457] that arithmetical truths are synthetic, here agreeing with Kant instead of Frege. This was possible for Russell because he thought, here diverging from Kant even more than Frege had, that logical truths are synthetic. Russell even hinted that all truths might be synthetic. Here Russell in a way anticipated Quine’s and Putnam’s position (though Russell’s point seems to have been not that the notion of analytic truth would be obscure, as Quine held, but that its extension was empty). Also the early Russell thought (see [Rus00, §12, page 24]) that though arithmetical truths are not analytical, they are yet necessary, a view that Quine would not have liked.

37Another example used by Kant were of course geometrical statements. These are even more controversial and difficult than arithmetical ones. Frege did not originally intend the thesis of logicism to apply to geometrical statements - in [Fre84, §14] (see [Fre50, §14, page 20] for an English translation) he holds that they are not empirical yet still synthetic, from which it follows that they must be synthetic a priori, just as Kant thought. However, later analytical philosophers extended it to them. Usually this is done by distinguishing between pure and applied geometry, where pure geometry is held to be analytical, while applied geometry is held to be part of physics and hence synthetic a posteriori. However, it can be argued that this kind of view suffers from the fallacy of equivocation; pure geometry is not geometry at all in the traditional sense, not what Euclid or any mathematician before the 20th century would have understood to be geometry. I think that it would be more
it is increasingly challenged by the radical empiricistic, "naturalistic" view according to which they are synthetic a posteriori. If the dominant view of analytical philosophers concerning arithmetic is correct, it shows at least that analytical judgments can be less trivial than Kant thought, and there are equally difficult epistemological questions concerning their possibility as concerning the possibility of synthetic a priori truths. Furthermore, if the view is correct this of course raises the suspicion that Kant is likely to have been wrong about the source and kind of metaphysical cognition in the same way as he was about the source and kind of arithmetical cognition. Also modern philosophers hold some metaphysical principles to be analytical; for instance, some mild deflationists such as Bob Hale and Crispin Wright think that some crucial metaphysical principles such as abstraction principles are analytical a priori.

There is also a minority of philosophers, including for example realistic phenomenologists such as Barry Smith in [Smi96], who still think that there are synthetic a priori truths and many of them think that many metaphysical truths are such; however, such philosophers commonly disagree with Kant both about the extent and more importantly about the basis for a priori knowledge. Husserl was convinced by Frege that arithmetical knowledge was logical and hence analytic and realistic phenomenologists generally follow him. Rather, they (and most discussion in today’s philosophy) would take such statements as "The same surface cannot be simulta-

38 Matters are complicated by the fact that phenomenologists often claim that such truths are based on experience in a wider sense, namely on a special kind of experience, phenomenological experience or experience of essences, Husserl's Wesensschau, so in a sense they are synthetic a posteriori. Also it is often claimed that such truths are analytical in a wide sense, wider than the one captured by Frege's definition. Carnap's theory of meaning postulates in [Car52] captures in a way both these truths and those which are analytic according to Frege’s definition. So the difference between this view and the first and second views is not as great as it might at first appear. I will later in Section 6.4.1 present an epistemological theory of Reinhardt Grossmann which represents a very interesting combination of the first and the third view.


neously both red all over and green all over.” as relatively uncontroversial (if not in themselves philosophically interesting) examples of synthetic a priori truths. As I will show in more detail later, phenomenologists generally distinguish between ontology and metaphysics and think that ontological knowledge is a priori while metaphysical knowledge is a posteriori, and they also distinguish between formal ontology whose truths are analytical and material ontologies whose truths are synthetic. More importantly, realistic phenomenologists think that synthetic a priori truths are grounded in objective relations between essences (of which we can according to them have intuitions, contrary to Kant’s view) rather than in subjective forms of sensibility and understanding as Kant thought. This view can be made more plausible by considering the several problems that have been found in Kant’s transcendental theory of synthetic a priori knowledge. Of course, this view requires realism concerning universals (as the essences of phenomenologists are just what scholastics called universals) while Kant’s view presupposes nominalism. However, Kant never justified the nominalism he presupposed, apparently thinking it obvious or already proved by others, while the modern realistic phenomenologists generally have arguments for their realism and against nominalism (taking account of the modern arguments in favour of nominalism by British empiricists and continental rationalists on which Kant presumably relied) so this is scarcely a disadvantage for phenomenologists. Furthermore many of these arguments do not presuppose the existence of synthetic a priori truths and are accepted also by philosophers who do not accept the existence of synthetic truths a priori - e. g. Quine at one time accepted the existence of universals though he rejected the whole distinction between analytical and synthetic truths - so such reliance on universals is not question-begging.

The second and third view have only rarely met in discussion. One example of debate between them occurred when Husserl and Schlick had a brief debate about a priori truths - Schlick criticized Husserl and other phenomenologists such as Max Scheler in [Sch18, 120-121] and in [Sch32a] (translated into English in’ [Sch49a] and Husserl responded in the Logical Investigations - but this debate ended in mutual misunderstandings. Many modern commentators such as M. M. Van de Pitte (in [VdP84]), Jim Shelton, Roberta Lanfredini and Paul Livingston (in [Liv02]) have discussed this debate at length. This has as might be expected resulted in a prolongation
of the debate, as the commentators do not agree even on the interpretation of Schlick’s or Husserl’s theories or on the relations between them, much less on whose arguments were more correct.

This debate strangely concentrated not on the existence of synthetic a priori truths but on that of material a priori truths, which one would have thought to be less controversial. It is generally thought today that if any truths at all are analytical then formal truths are analytical, and indeed the least controversial examples of analytical truths, but not all purportedly analytical truths are formal (see e. g. [Car52, page 66]). There have been a few dissenters; Jerrold Katz thought (see e. g. [Kat72, page 119]) that logical truths are not analytical truths in the narrow sense of analyticity he distinguishes. Even such utterly trivial and relatively uncontroversial analytical truths as ”Bachelors are unmarried.” are in one obvious sense material truths, since they depend on the interpretation of the non-logical constants that occur in them. If you interpret ”bachelor” as referring to soldiers but continue to interpret ”unmarried” in the ordinary way the sentence does not continue to be true, and certainly if you reinterpret ”green”

---

39 Austin’s English translation of Frege’s definition, which directly led to the concept of analyticity dominant in analytical philosophy, was a bit syntactically ambiguous on this score, which may have sowed some confusion. When Frege said according to Austin that a proposition is analytical if in following it right back to primitive truths we ”come only on general logical laws and definitions”, it is not clear whether only the laws or also the definitions are also supposed to be general and logical, i. e. to involve only logical i. e. formal concepts. Apparently because of this ambiguity, Frege has been interpreted (e. g. by Guillermo Haddock in [Had08, pages 132,133]) as holding that all analytical truths would have to be formal. Haddock claims that unlike Carnap, Frege would not have held such a statement as ”All bachelors are unmarried.” to be analytic. Frege’s following definition of syntheticity is also a bit obscure (and it is not clear that it is the contradictory of analyticity) which adds to the confusion. However, the German original of the definition of analyticity

nur auf die allgemeinen logischen Gesetze und auf definitionen

makes it clear that the qualifications ”general” (”allgemein”) and ”logical” (”logisch”) were meant to apply only to the laws, not to the definitions. Therefore the definitions could contain non-logical concepts and so Frege probably meant his definition in the way analyticity was usually later understood in analytical philosophy, so that he implicitly agreed with Carnap that ”All bachelors are unmarried.” was analytic, contrary to Haddock.

40 There are problems with Katz’s arguments for his view. Katz appealed to Kant’s definition of analyticity (though this was not his only premise; he also appealed to linguistic intuitions). It may be true that Kant’s definition of analytical truths could be understood so that logical truths would not be analytical. However, as we have seen, clearly Kant himself thought that logical truths were analytical, so he did not intend his own definition to be understood in the way Katz did. Therefore Katz’s appeal to Kant does not really support his theory.
as referring to the property crimson or crimson things the sentence "The
same surface cannot be simultaneously both red all over and green all over."
does not continue to be true. However, Schlick apparently thought that not
only such truths but even the truths that the phenomenologists thought to
be synthetic a priori are not only analytic but even formal, saying in [Sch49a,
page 285]:

The error committed by the proponents of the factual a priori
can be understood as arising from the fact that it was not
clearly realized that such concepts as those of colors have a for-
mal structure just as numbers do or spatial concepts, and that
this structure determines their meaning without remainder.

A charitable interpreter might suggest that Schlick was using the word "for-
mal" in a wider sense than we usually do today, though in this case it be-
comes rather unclear what the concept of formality he employed was. How-
ever, it is also possible that Schlick was just confusing the distinct notions
of formality and analyticity, and if this is correct then surely this invalidates
his criticism of Husserl. This of course does not imply that Husserl’s theory
would be correct or even meaningful, for there are many serious obscurities
in Husserl’s theory which cry out for a careful criticism by the methods of
analytical philosophy, but Schlick did not succeed in giving such a criticism.

It may then be that we are here dealing with different notions of form,
and indeed van de Pitte distinguishes three senses of the word "formal"
in [VdP84, page 201], distinguishing forms of experience in Kant’s sense
from the standard sense in which only propositions of pure logic are formal
and from the sense of the word Schlick is using here. However, van de Pitte
muddles matters up further in [VdP84, page 202] by claiming that phenome-
nologists would have admitted what Schlick says they did not realize. van de
Pitte claims that nothing is more central to phenomenology than the insight
that every concept has a formal structure, and that when Husserl asserts
that some a priori propositions have material content he means that they
present a structural law. However, while the first of van de Pitte’s claims is
probably correct, the second is clearly false, so Schlick did not misrepresent
the relationship between his view and the view of the phenomenologists so
much as van de Pitte claims. Husserl did not employ the word "material"
so idiosyncratically that he would have meant something structural with it.
Van de Pitte is correct that phenomenologists would have admitted that such concepts as colours have a formal structure, but not that they would have accepted that it would determine their meaning without remainder; these are two entirely different claims to which they would in my view have had different responses. Furthermore, the latter claim is extremely implausible; it is based on Schlick’s structuralism, which as I will show at detail in Section 3.1.3 of this dissertation, has been decisively refuted by Newman in [New28]. Because of this I think that even if the phenomenologists were wrong, which seems quite possible though far from certain to me, their error was not where Schlick thought it to be; the truths that they thought synthetic a priori may be synthetic a posteriori or such analytical a priori truths as are not logical in the narrows sense and hence material, but certainly not formal or structural.

Since Quine’s criticism of the distinction between analytic and synthetic

\[41\] At least some of Schlick’s criticisms of phenomenologists are unfair. Schlick claims (see [Sch49a, pages 277-278]) that the phenomenologists give a new meaning to the term "experience", and so commit a fallacy of equivocation. It is indeed true that phenomenologists have a different conception of experience than e. g. Kant (whom they criticize) did, but this need not imply that they use the word "experience" in a different sense. Of course, early logical positivists like Schlick also had a very different conception of experience than Kant, since they held (see e. g. [Sch36, page 359] and [Car67, §65, pages 103-106]) that original experience is without a subject, while Kant would have thought it true, probably even analytically true, that experience must have a subject who experiences (though this becomes typically complicated with Kant, whose theory distinguishes different subjects, both a phenomenal subject and a noumenal one, and according to some commentators even a third transcendental subject). On the other hand, Otto Neurath named a perceiver thrice in his peculiar formulation of protocol sentences (see [Neu32, page 207] and [Neu59, page 202]). Since protocol sentences were supposed to be about the verifying experience (as I will explain in more detail later), translating his view into the "material mode" we see that Neurath thought that experience had a subject, even if he tried to physicalize this subject which had earlier been viewed as immaterial. We then see that logical positivists were not unanimous about this view of experience. Logical positivists also generally thought that original experience was not constituted by mental acts either, while earlier analytical philosophers has thought that while such experience was without a primitive subject it did contain primitive mental acts. Husserl’s later concept of experience is in this respect closer to Kant’s, since he also held that experience had a subject (and indeed distinguished several subjects as Kant did, though not exactly in the same way); however, Husserl’s view changed, and his earlier view which held the subject to be constituted by mental acts was like that of earlier analytical philosophers and hence a bit closer to Schlick’s). Therefore Schlick also might be accused of giving a new meaning to the word "experience". However, the word "experience" is so very indefinite in meaning that both can be held to just sharpen its ordinary meaning, although they do it in different ways. However, since such philosophical explication of the concept of experience easily leads to a conception of experience which differs from the ordinary one, it is by no means a trivial question whether either Schlick’s or Husserl’s (or indeed Kant’s) conception of experience is even coherent.
truths, analytic truths which are not logical have come to appear to many philosophers who still call themselves analytic philosophers as problematic as synthetic truths a priori were to the early analytic philosophers. This narrows the gap between the second and third views, especially given the fact that as I already argued, the notion of experience may be ambiguous and hence the distinction between the second and third view is not as sharp as it at first seems. Quine himself did not find the distinction between logical and non-logical truths very problematic (though he did vacillate in his opinions concerning this subject). However, many later analytic philosophers have found it also dubious. The question of what truths are logical and what not is not an easy one, and has been much discussed without much agreement being reached (for a good treatment of the topic see John MacFarlane’s exposition in \cite{Mac00}). Such philosophers as John Etchemendy indeed see (e.g. in \cite{Etc90}) this distinction as more problematic than the more general distinction between analytic and synthetic truths. Such modern extreme empiricists as Devitt do not find the question of how we know logical truths any easier to answer than the question of how we could know synthetic a priori truths.

Lately most analytic philosophers have come to think that the logicist project of Frege, Russell, Whitehead and Husserl has failed. The foundations they sought in higher-order logic have instead in the view of many mathematicians been found in set theory, but the epistemological and ontological status of set theory is unclear. Because of this there is little unanimity or clarity about what this would mean for general epistemology. Doubts regarding whether mathematical truths can be viewed as logical have led to doubts regarding whether they can be analytical truths at all. This has often been taken to support extreme empiricism with regard to mathematics. However, it could equally well be taken to suggest that we should re-evaluate the rejection of synthetic a priori truths, not only in Husserl’s sense but also in Kant’s sense (which of course would not imply that we should accept Kant’s explanation of our knowledge of synthetic a priori truths).

However, since as I showed there is no clarity or unanimity about where the boundaries of logic lie or how to draw them, it cannot be taken as a settled fact that logicism has failed. It cannot in my view even be taken as sure that set theory is not part of logic\footnote{I will argue later that set-theoretic membership can be viewed as a kind of predication.}. Indeed not all analytical philosophers

65
agree that the failure of logicism would be settled; Nino B. Cocchiarella protested against this view in [Coc87], and lately Bob Hale and Crispin Wright have tried to defend a Neo-Fregean logicist program in [HW01]; other authors such as Richard G. Heck (see [Hec11]) and Charles Parsons have also pointed out that at least parts of Frege’s logicist program could be salvaged. Therefore it would be premature to say that our epistemological research would have to start from the failure of logicism as a settled fact.

Even this brief discussion ought to show that the issues at play in the debate between the three epistemological positions are very complex and difficult, so difficult that even characterizing the relations between the different positions accurately is hard, never mind deciding correctly between them. Fortunately, we do not have to decide which of these three general anti-Kantian epistemological positions is correct here, though addressing them will become imperative in deeper metametaphysical investigations. Metametaphysics should address general epistemological problems only so far as they are relevant for metaphysics, and address them at each point of its argumentation only so far as they are relevant at that point.

I will touch them briefly later in this work, arguing that the most influential arguments for the extreme empiricist view, Quine’s arguments, are bad because they are dependent on verificationism, which has been shown to be very problematic, and their verificationist premises lead to semantic holism, which has unacceptable consequences. I argue that gives us some reasons to reject the extreme empiricist view, since Quine’s arguments have been most influential in establishing it. However, these reasons are not yet conclusive, since other less influential arguments for this conclusion (e. g. Devitt’s arguments) may not share all of the defects of Quine’s arguments (though Devitt also appeals to Quine, so if I am right that the premises of Quine’s argument are indefensible, this also weakens the case of Devitt).

However, at this stage of the argument the vital point to note is that

---

Predication, however, is generally considered a logical concept, so this would indicate that at least the basic concepts of set theory are logical. It is interesting to note that when we delve into the origins of modern logic and mathematics, we see that Giuseppe Peano held set-theoretical notions to be logical (see e. g. his listing of his signs in [Peano, page vi] where he divides all signs he uses into logical and arithmetical and counts all set-theoretic signs as logical), though he was not himself a logicist. If Peano’s understanding of what notions are logical were combined with any of the now popular reductions of arithmetical notions to set-theoretical ones, whether that of Zf or less standard ones like Quine’s NF, NFU and ML, we would get a new version of logicism.

66
if any of these three general epistemological theories is correct then Kant’s premises for his view that we cannot have knowledge of things in themselves are incorrect. Therefore the fact that philosophers today generally hold some one of them and have plausible arguments for it suffices to make Kant’s original reasons for rejecting metaphysics in the traditional sense (what is sometimes called transcendent or speculative metaphysics) irrelevant and outdated.

Because of this Kant’s prolegomena do not work any more as prolegomena to any future metaphysics that will be able to come forward as a science; indeed, they did not work very well as such prolegomena even in Kant’s own time, since as I have shown Kant misinterpreted the traditional metaphysics he was criticizing. Nevertheless, as I have already argued Kant was right in thinking that metaphysics needs prolegomena, so we need to find new prolegomena. Of course, such prolegomena are not entirely new today, since later philosophers have supplied many promising candidates for at least the beginning of such prolegomena, such as Quine’s theory of ontological commitment, at least as it is developed by Peter van Inwagen; I will consider them in this dissertation and try to develop further.

\[43\] There are of course still some philosophers who cling to some of Kant’s ideas. However, most of them, such as internal realists of Hilary Putnam’s kind or Philström’s pragmatic realism, which are often supposed to be analogous to Kant’s empirical realism combined with transcendental idealism, only take very general inspiration from Kant. They do not generally take Kant’s original problem, that of synthetic a priori judgments, as their main point of departure. Therefore any arguments they would have against metaphysics or for a specific kind of metaphysics should have to be examined on their own terms. In fact their premises are closer to the verificationism of the logical positivists, which I will discuss in the next section of this work, than to Kant’s original premises, so that my criticism of verificationism to some extent also applies to their arguments. In any case, Putnam at least does not seem to completely reject metaphysics, but like Kant he rejects it in some sense and accepts it in another, weaker sense, and Philström quite explicitly accepts metaphysics.
Chapter 3

Verificationism and Other Metaontological Principles in Logical Positivists

3.1 Different Elements in Logical Positivism: Empiricism, Structuralism, Foundationalism, Coherentism and Syntacticism

3.1.1 Empiricism and Verificationism in Logical Positivism

Of course, though modern philosophers do not usually accept Kant’s premises for the rejection of traditional metaphysics this is far from implying that philosophers would generally admit the possibility of transcendent or even of immanent metaphysics. There have been many other reasons (incompatible with each other) why modern philosophers who make use of a critical method hold as negative opinions about metaphysics as Kant did or even more negative ones. While these later attacks reject most of Kant’s premises, Kant’s misunderstanding of the genuine nature of metaphysical research continues to influence them (as I will show); many later critics of metaphysics assume uncritically on the basis of Kant’s ideas that metaphysics would have to be synthetic a priori, or at least that it would have to be synthetic, and therefore that arguments against the possibility of synthetic knowledge a priori would also be automatically arguments against the possibility of metaphysics.
Empiricism is perhaps most often presented as an obstacle to metaphysics, even though the original inventor of metaphysics as a separate discipline, Aristotle, was, as we have seen, an empiricist (at least in a weak and possibly even in a rather strong sense). This is perhaps partly because of the continuing effect of Kant’s misunderstanding of metaphysics as synthetic a priori by definition, since empiricism is disdaining of any kind of a priori knowledge and is blatantly incompatible with the existence of synthetic a priori knowledge. However, it is also partly because empiricism has been taken to radical extremes which might threaten even a conception of metaphysics which does not take its claims to be synthetic and a priori. Such extreme forms of empiricism are indeed obstacles to metaphysics but also to a lot else that even opponents of metaphysics would price.

A rather extreme kind of empiricism (in some ways slightly less extreme than the later kinds of empiricism of Quine or Devitt, but in other ways still more extreme) represented by such philosophers as Auguste Comte (1798-1857), Ernst Mach (1838-1916) and Richard Avenarius (1843-1896), was called positivism. In the beginning of the twentieth century this was combined with the methods of modern logic resulting in a sophisticated version of positivism called logical positivism or logical empiricism. Logical positivism was one form of analytical philosophy; sometimes the anti-metaphysical orientation of logical positivism is ascribed to all analytical philosophy, but this is utterly erroneous, since the earliest analytical philosophers were not opposed to all metaphysics. In fact the recent movement of analytical

---

1 I will be mostly using the term "logical positivism" instead of "logical empiricism" because I think that it is a better name for the movement in question. As I argued there is nothing in the word "empiricism" to suggest opposition to metaphysics, while the word "positivism" has always signified some kind of opposition to metaphysics (although the way in which different positivistic theories are opposed to metaphysics has been very different), which was a very prominent characteristic of the movement. Logical empiricism might be understood so that it would be a broader movement than logical positivism, possibly including for example middle-period and late Russell (who was not a logical positivist), and if so I am not opposing logical empiricism (in such a broad sense) but only logical positivism.

2 Nino B. Cocchiarella distinguishes in [Coc87, pages 1-2] three stages in the development of analytical philosophy, two metaphysical stages and one anti-metaphysical stage. The two metaphysical stages are the stage of "extreme" realism and the stage of logical atomism. Much of the new analytical metaphysics such as Armstrong’s famous metaphysics can be viewed as a return to the second metaphysical stage. As I have argued in Section 2.1.4 that the fundamentalism which is part of logical atomism is not well justified, I think that we should rather return to the very first stage, or try to find some intermediary position between the doctrines of the first and second stage.
metaphysics is a genuine continuation of the metaphysical trends of original analytical philosophy. Indeed, all the founders of analytical philosophy were arguably not empiricists at all - e.g. it is highly doubtful if Frege could be called an empiricist in any non-vacuous sense, since as we have seen he accepted not only a priori knowledge but even synthetic a priori knowledge in the case of geometry - much less extreme empiricists of the kind logical positivists were. Logical positivists mostly represented (along with most other analytical philosophers) the second of the three anti-Kantian epistemological traditions I listed in the previous section of this work, though as I will show they were often inconsistent with this. One of the most extreme kinds of empiricism ever propounded was encapsulated in the famous verificationist criterion or principle of the logical positivists. According to a rather early and obscure but common formulation of the principle every meaningful (synthetic) statement had to be verifiable by means of experience and logical positivists thought that metaphysical statements were not verifiable (and were synthetic), so this gave them an apparent reason to reject metaphysics.

This reason for rejecting metaphysics should be as outdated as Kant’s epistemology. For a long time it seemed that logical positivism had been rejected by the philosophical community. John Arthur Passmore declared already in 1967 in [Pas67] that logical positivism was as dead as a philosophical movement ever becomes. Unfortunately, that turned out to be not very dead at all3. Huw Price has called metaphysics after Carnap the ghost who walks, but I will argue that in truth is is the logical positivism, of which Carnap was the most prominent representatives, that is the ghost who walks. In fact logical positivist conceptions lie hidden - whether implicitly affirmed or equally implicitly denied - at the background of several modern discussions about truth and about realism and anti-realism within analytical philosophy, so that it is hardly possible to even properly understand the modern debates without having some knowledge of the history of logical positivism.

Because of this I will have to spend some time on discussing logical positivism, more than I spent discussing Kant’s more severely outdated the-

---

3Passmore also said that logical positivism considered as a doctrine of a sect, has disintegrated, and this may indeed remain true. However, it still lives, as I will show later at more length, as a doctrine held by philosophers that do not belong to any common sect. I will argue, however, that it should die also as a doctrine because of the many very serious problems that have been repeatedly pointed out in it.
ories. However, what I say in this section of my work will mostly not be original, though going through it seems to me necessary so that I may dismiss considerations of verifiability from my later work; I will gather from the literature what seem to me the most important reasons because of which the main doctrines of logical positivism are rejected by many philosophers today. Those readers who are already fully convinced that logical positivism was fundamentally mistaken might skip this whole chapter of my dissertation.

### 3.1.2 A Proposed Re-evaluation of Logical Positivism

Recently a new interpretation and evaluation of logical empiricism has become popular, which sees it as derived from the Marburg school of Neo-Kantian philosophy and Poincare’s conventionalism (expressed in such sources as Henri Poincare’s [Poi05]) rather than empiricist sources. One source of this new interpretation is in [Fri87], where Michael Friedman gives a new interpretation of Carnap’s Aufbau, and Friedman has gone on to extend such an interpretation to all logical positivism in [Fri91] and expanded this into book form in [Fri99]4. Other proponents of this kind of interpretation are Alan W. Richardson (e.g. in [Ric90] and in [Ric97]) and Alberto Coffa, Thomas Ricketts etc. Bryan G. Norton also defended a similar view about Carnap’s philosophy earlier than either Friedman or Richardson in [Nor77]. Friedman de-emphasizes the importance of the verification principle in Carnap’s philosophy and instead sees the structuralism and conventionalism also presented in Aufbau as central.

While Friedman, Richardson and others are surely correct that there are Kantian and structuralist and conventionalist elements in the thinking of logical positivists, yet it is also quite clear that there are also elements of extreme empiricism in them. Indeed, they never go so far as to explicitly deny that those elements exist5, only downplay their importance. Friedman

---

4Friedman says in [Fri91, page 505] that no longer challenged by logical positivism as a live philosophical option, we can consider it impartially from a historical point of view. However, the main ideas of logical positivism are very much alive, as I will argue, so it cannot yet be considered from a purely historical point of view, but must be treated as a live philosophical option. Indeed, Friedman’s suggestion fits badly with his own assertion (in [Fri91, page 519]) of parallels between the positivists’ situation and our own post-positivist situation. We could only consider logical positivism impartially when (if ever) our philosophical situation differs entirely from the positivists’ situation.

5Indeed, in [Fri87, page 536], Friedman admits that that there is no doubt at all that the Aufbau does defend empiricism and phenomenalism.
says in [Fri91, page 510] that the positivists’ main philosophical concerns did not rise within the empiricist tradition at all. Even if this were true, it would allow that logical positivists were empiricists (and even that some of their philosophical concerns rose from the empiricist tradition). So this kind of interpretation seems to differ from the traditional one only in what it emphasizes. Surely it does not suffice for a genuine re-evaluation of a thinker to just propose a new view of his concerns; the value of a thinker’s work is based on whether the assertions he makes are correct or justified (and possess other values such as informativeness, fruitfulness, etc.). If a philosopher makes a blatantly incorrect assertion, it is no defence that he made it inadvertently.

However, it does not seem to be true. Friedman and Richardson contrast Carnap and other logical positivists with Bertrand Russell, whose External World program they take to represent a more traditional variety of empiricism. However, nearly all the dissimilarities they point out between the concerns of the logical positivists and Russell are based on either ignoring or misrepresenting Russell’s philosophical development.

Friedman says in [Fri91, page 150] that the initial impetus for the positivists’ philosophizing came from late nineteenth-century work on the foundations of geometry. However, if we look at Russell’s earliest (theoretical rather than political) book [Rus97], we see that at least part of Russell’s initial impetus for philosophizing also came from the very same work on the foundations of geometry. Also Friedman admits that many of the very geometers that gave the logical positivists their impetus held an empiricist conception of geometry, which implies that they were influenced by the empiricist tradition, so Friedman’s attempt to separate the motivations of the logical positivists from the empiricist tradition is futile.

Richardson sees in [Ric90, page 12] the difference between Carnap and Russell in the fact that the over-all program of the Aufbau is more clearly directed toward Kantian issues and problems than those of traditional empiricism. Richardson thinks that Carnap was mostly interested in the objectivity of science and sought to ground this objectivity in the structural properties of the objects of science. However, Russell himself had been strongly influenced by Kantian, namely Neo-Kantian (and even Hegelian) idealist philosophy in his earliest philosophy (as can also be seen in the already mentioned book [Rus97]). Kantian issues and problems surely remained im-
portant for him also in his whole philosophical development, even if he later sought solutions for those problems from more empiricist doctrines, having grown dissatisfied with Kantian doctrines (as the logical positivists also did). Nor does structuralism serve to establish any significant difference between Russell and the logical positivists; there were structuralist elements also in the last stages of Russell’s External World program, as famously seen in [Rus27], and Russell also sought to ground the objectivity of science in these structural properties. Therefore even if structuralism were more defensible than empiricism, this would not give logical positivists any superiority over Russell. However, if a famous argument by Newman against structuralism, which I will present in Section 3.1.3, is correct, then both Carnap and Russell were speeding down a blind alley in their structuralism, so that objectivity must be grounded otherwise if it can be grounded at all. So Friedman’s and Richardson’s attempt to draw a major difference between Carnap’s and Russell’s concerns and issues just does not work.

While there are then no great differences in Russell’s and the logical positivists’ concerns, there were in their eventual theories. Friedman and Richardson argue that the logical positivists were not naive empiricists. This may be so; however, Russell surely was not a naive empiricist either; indeed, it can be questioned in what sense he was an empiricist at all, as he seems to have at least at some times admitted the existence of synthetic a priori knowledge. However, I will show that Carnap and other logical positivists committed themselves (at least in the early stages of their development) to a far stronger form of empiricism than Russell ever did, namely to verificationism. Surely then if such a stronger form of empiricism is less defensible than Russell’s weaker one (as I will argue it is), this is a reason to evaluate them more negatively than Russell, even if empiricism were less important for them than for Russell as Friedman and Richardson claim. Another clear difference is that Russell steered clear of the extreme conventionalism of some logical positivists; however, I will argue that such extreme conventionalism is quite indefensible, so this makes Russell’s theory better.

In fact the presence of Kantian, structuralist, conventionalist and coherentist doctrines in logical positivists (and in Russell) gives us reasons for evaluating that movement (and indeed Russell too, though in a lesser degree) more negatively, rather than more positively, since not only are there as strong arguments against those doctrines as against verificationism, but
it also poses the risk of radical incoherence in logical positivism (though it poses the risk of incoherence also in Russell’s [Rus27]). Any kind of strong structuralism or conventionalism or coherentism or Kantianism is arguably incompatible with even kinds of empiricism far weaker than verificationism, and logical positivists did generally formulate structuralism, conventionalism and coherentism in a very strong way. If consistently carried out, such a strong structuralism or conventionalism or coherentism rejects entirely the idea that rational beliefs or even concepts have to be somehow connected to experience.

The relationship of logical positivists to empiricism was then very problematic from the first and arguably blatantly inconsistent; an inconsistency which I will argue remains in much of current philosophy influenced by logical positivists. Clearly if the epistemological doctrines championed by the logical positivists which they took as premises in their critiques of metaphysics are incompatible with each other, then the criticisms based on such incompatible premises cannot all be correct. However, it is possible that none of them is correct, and I will argue that this is in fact the case since the premises of none of them are tenable if interpreted strongly enough that they genuinely imply the illegitimacy of (even transcendent) metaphysics.

While Friedman does not explicitly deny that the logical positivists were empiricists he does deny explicitly (e.g. in [Fri91, page 506]) that they would have been foundationalists. This is an incoherent position, since empiricism as well as rationalism in the ordinary senses of the words are both foundationalistic theories, competing theories about the foundations of knowledge which both assume knowledge must have. The rejection of foundationalism would then require the rejection of both. However, Friedman misunderstands what foundationalism is usually taken to be. He supposes (in [Fri91, page 508]) that it would imply that philosophy as a discipline is foundational with respect to the special sciences. Of course, this is not what foundationalism in the usual sense of the word means or implies at all (though some foundationalists have indeed held such a view, which is not completely without reasons⁶). It is the experiences of scientists themselves

⁶In fact if we view scientific knowledge historically, as Kuhn and Friedman urge us to do, then it is clear that philosophy has in fact influenced the development of natural science (as has pure mathematics). In fact the natural sciences have emerged from philosophy. This gives us a reason to say that philosophy has been partially foundational to natural science (though of course not that it would give natural science all of its principles,
when performing their experiments that an empiricist foundationalist sup-
poses to be (at least a significant part of) the foundation to the scientific
knowledge they produce. This is quite compatible with the claim Friedman
ascribes to the logical positivists that philosophy must follow the evolution
of special sciences so as to test itself\footnote{Some logical positivists, however, may have been guilty of the incoherent position Friedman inadvertently ascribes to them, though definitely not all. Schlick was very clearly a foundationalist (as seen in \cite{Sch34}) - Friedman of course does not mention this article in \cite{Fri91}, as it does not fit at all to his view of logical positivism - and so was at least the early Carnap. However, it can be argued that if the epistemology Otto Neurath developed in \cite{Neu32} (translated in \cite{Neu59}) and in \cite{Neu34} (and Hempel defended in \cite{Hem00}) was epistemologically coherentist, as Schlick argued in \cite[page 84]{Sch34} (translated in \cite[page 214]{Sch59}) and as Bertrand Russell already argued in \cite[page 140]{Rus40}, it does not count as a genuine version of empiricism at all (as Russell also concluded in \cite[page 148]{Rus40}), though Neurath himself frequently called himself an empiricist. However, as verificationism is a kind of empiricism, Neurath’s theory would not count as genuinely verificationistic either. However, Neurath’s stressing of the importance of protocol sentences would seem to indicate some remnants of empiristic foundationalism in his philosophy, since (as I will show later) protocol sentences were conceived as as sentences speaking of experience. A consistent and complete rejection of foundationalism would instead require that - as Donald Davidson proposed in \cite[page 313]{Dav86} - we should give up the distinction between observation sentences and other sentences. These remnants of foundationalism in Neurath’s philosophy are all that prevented it from lapsing into a complete relativism.}.

While Friedman, Richardson and others are correct that such doctrines
distinct from empiricism as structuralism, coherentism and conventional-
ism must be taken into account in presenting and evaluating the theory of
logical positivists, I will argue in the next subsections of this chapter that
including such elements does not improve the prospects of logical positivism
in the slightest. At most Friedman’s and Richardson’s arguments justify
a partial \textit{re-interpretation} of logical positivism, but not a \textit{re-evaluation} of
it. I will argue that even if we could just ignore the empiricist elements
in logical positivism, this would not give us any reason to evaluate logical
positivism any more favourably, for structuralism, coherentism and conven-
tionalism are faced with even more serious problems than verificationism.

\footnote{Some logical positivists, however, may have been guilty of the incoherent position Friedman inadvertently ascribes to them, though definitely not all. Schlick was very clearly a foundationalist (as seen in \cite{Sch34}) - Friedman of course does not mention this article in \cite{Fri91}, as it does not fit at all to his view of logical positivism - and so was at least the early Carnap. However, it can be argued that if the epistemology Otto Neurath developed in \cite{Neu32} (translated in \cite{Neu59}) and in \cite{Neu34} (and Hempel defended in \cite{Hem00}) was epistemologically coherentist, as Schlick argued in \cite[page 84]{Sch34} (translated in \cite[page 214]{Sch59}) and as Bertrand Russell already argued in \cite[page 140]{Rus40}, it does not count as a genuine version of empiricism at all (as Russell also concluded in \cite[page 148]{Rus40}), though Neurath himself frequently called himself an empiricist. However, as verificationism is a kind of empiricism, Neurath’s theory would not count as genuinely verificationistic either. However, Neurath’s stressing of the importance of protocol sentences would seem to indicate some remnants of empiristic foundationalism in his philosophy, since (as I will show later) protocol sentences were conceived as as sentences speaking of experience. A consistent and complete rejection of foundationalism would instead require that - as Donald Davidson proposed in \cite[page 313]{Dav86} - we should give up the distinction between observation sentences and other sentences. These remnants of foundationalism in Neurath’s philosophy are all that prevented it from lapsing into a complete relativism.}
These problems have all been pointed out before, yet philosophers are not generally sufficiently aware of them, so I must discuss them here before moving on. Strangely Friedman himself has in other writings shown that he is aware of these problems. If the arguments against structuralism, conventionalism and coherentism I will present are valid, we cannot base any valid argument against the possibility of metaphysics on structuralism, conventionalism or conventionalism. After discussing these three theories I will return to verificationism, which presents a more serious menace for metaphysics.

3.1.3 Structuralism and Structural Realism in Logical Positivism and Outside It

Ironically, structuralism\(^8\) is often presented in current discussions in the philosophy of science as a form of scientific realism, structural realism, as was famously done by John Worrall in [Wor89] and by Grover Maxwell in [Max70]. However, structuralism has also been connected with doctrines which are as far as possible from realism. Though Barry Gower initially in [Gow09, page 74] counts the Marburg Neo-Kantian Ernst Cassirer as a structural realist, he later reveals in [Gow09, page 90] that Cassirer had in fact viewed structuralism as connected with idealism. Logical empiricists defending structuralism generally held - partly on its basis - the very notion of realism (as well as idealism) to be meaningless\(^9\). The Neo-Kantians like Cassirer...

\(^8\)In this section of my work I will be dealing with structuralism as a theory in general philosophy of science. I will not be dealing with structuralism in the philosophy of mathematics, which is a separate view. Structuralism as a view concerning mathematics is on the face of it more plausible than concerning other sciences like physics, since mathematical objects can quite plausibly be argued to be structures (abstract structures in a sense defined below). Nevertheless, even in their case we must distinguish between the structures they are and the (higher order) structures they themselves have, a distinction which mathematical structuralism often tends to blur.

\(^9\)It is likely, however, that (as Gower argues in [Gow09]) Schlick at first held a structural realist position, but this was before he became a logical positivist (e. g. in [Sch18]). When he converted to logical positivism (under Wittgenstein’s influence) he rejected structural realism since he completely rejected realism, even restricted realism, as a metaphysical thesis (as can be clearly seen in [Sch49b, page 107]), but he probably did not reject structuralism (though [Sch49b, page 107] is not really with consistent with structuralism). Some of Schlick’s writings such as [Sch26] (translated in [Sch79a]) are transitional as Schlick is already in them moving toward positivism, rejecting metaphysics strongly, but yet clinging to structural realism, and strengthening his form of structuralism in a way I will show is unfortunate. Though Gower is then correct in claiming that Schlick was a scientific and structural realist, he seems to go astray in going on to claim (in [Gow09, page 100]) that Schlick remained a scientific realist, though admittedly the claim may be slightly ambiguous. Schlick did continue to call himself an empirical realist (as Kant had...
sirer and logical positivists can then more appropriately be called structural anti-realists rather than structural realists. Also recently some anti-realist philosophers such as van Fraassen have accepted a form of structuralism. Such philosophers of science as Stathis Psillos have argued in articles such as [Psi06] that structural realism is not the best form of scientific realism, if it is a form of realism at all. While Psillos attacks structural realism from the point of view of a stronger form of realism, Friedman himself is not happy with the notion of structural realism for an opposite reason, since he holds with the logical positivists that realism is a metaphysical theory and that metaphysical theories are meaningless.

At least what is today called ontic structural realism, a position held such philosophers as Steven French James Ladyman and Don Ross which holds that all that exists are structures, i.e. mathematical entities, and that physical entities are identical with them or reducible to them, is nothing else than Pythagoreanism in a modern guise. Rather than being an anti-metaphysical theory, as logical positivists hoped their version of structuralism to be, is a very extravagant metaphysical theory. James Ladyman and Don Ross at least admit that metaphysics is meaningful and possible and that his ontic structural realism is a metaphysical theory and that logical positivism was in great part mistaken. Though their brand of rather extremely naturalized scientific metaphysics as presented is very different methodologically from traditional metaphysics, it yet clearly counts as metaphysics (and as metaphysics in a strong sense, speculative metaphysics).

Worrall sought to find a middle way between forms of anti-realism such as phenomenalism and constructive empiricism and ordinary forms of scientific realism, a kind of theory called restricted realism or selective realism. Worrall thinks that we can know that the statements made by a scientific theory about the structural i.e. formal properties of theoretical i.e. unobservable entities are literally true but other statements made about theoretical entities can not be known to be true. Worrall thinks that the ‘no miracles’ argument is effective against anti-realism; if scientific theories were not literally true at least in their structure then their empirical success would be called himself, though not in quite the same sense), but this is not the same thing as scientific realism, which (as the word is generally used today) refers to the metaphysical realism Schlick rejected. Schlick’s empirical realism would not usually be taken as a form of scientific realism, any more than Kant’s empirical realism or Putnam’s rather similar internal realism.
miraculous. However, Worrall cannot yet accept ordinary (unrestricted) scientific realism because of another argument. Worrall’s argument against a form of scientific realism stronger than structural realism in [Wor89] is based on pessimistic induction; many anti-realists (most famously Larry Laudan in [Lau81], though Laudan did not yet use the phrase "pessimistic induction") have used such an argument against any form of scientific realism, even restricted or selective realism. That the theoretical content of previous theories has so often been shown to be false by later theories gives inductive support to the claim that the theoretical content of any theory so far as it goes beyond structural statements would be false.

However, it seems to me that pessimistic induction could just as well be directed against the view that scientific theories are even empirically adequate. Scientific theories that have been abandoned have usually had empirical consequences that were false - after all, this is usually just why they were abandoned. It must be remembered that the common notion of empirical adequacy is rather strong. E. g. van Fraassen says in [vF80, page 12] that a theory is empirically adequate exactly if what it says about the observable things and events in the world is true. A theory need then not be empirically adequate just if what is says about things that have in fact already been observed is true. Worrall claims (see [Wor89, page 109]) that the development of science is essentially cumulative at the observational level but not at the theoretical level. However, we need to challenge the assumption that the development would be essentially cumulative even at the observational level. Worrall takes the change from Newton’s theory to Einstein’s as an example of a change which is cumulative at the empirical but not the theoretical level. He says:

But for a whole range of cases (those cases, of course, in which the velocities involved are fairly small compared to the velocity of light) the predications of the theories will be strictly different but observationally indistinguishable.

However, this hardly suffices for the claim that the development here would be essentially cumulative. If the observer moves at velocities close to the speed of light (which is not only logically but even nomologically possible; it is not unlikely that future astronauts will do so), then the predictions of the theories will be clearly observationally distinguishable. This already implies
that Newton’s theory is not even empirically adequate from the point of view of Einstein’s theory in the very strong sense of the phrase ”empirically adequate” used by van Fraassen (and many others). Of course Aristotle’s and Ptolemy’s or Descartes’s physical theories were not then empirically adequate either. What pessimistic induction then shows, if it is at all a valid form of inference\textsuperscript{10}, is that (current and future) scientific theories are at most only approximately empirically adequate. Worrall argues that there are difficulties in the concept of approximate truth; however, the anti-realist who is not a complete sceptic (like a constructive empiricist such as van Fraassen) must face the very same difficulties. Therefore he should take seriously the attempt of philosophers like Ilkka Niiniluoto in [Nii87] and Graham Oddie in [Odd86] to develop detailed theories of approximate truth and truthlikeness. The technical difficulties, which are indeed enormous (as the above books reveal), are not significantly easier if theories are restricted to theories speaking solely about observable entities. If, however, these difficulties can be overcome, or the whole pessimistic induction argument is fallacious, then there is no obstacle for claiming that we can have approximately true beliefs also about the non-structural properties of unobservable entities. Because of these reasons Worrall’s argumentation against full-blown scientific realism does not convince me.

The basic problem with structuralism was already noticed by M. H. A. Newman in [New28], when critically examining Bertrand Russell’s causal theory of perception\textsuperscript{11}. What is strange is that Friedman himself has admitted, in [DF89], an article written together with William Demopoulos, that Newman’s argument poses a very serious problem for structuralism, and has also admitted (in [DF89, pages 195-197]) that it also applies to the kind of structuralism held by Carnap in the Aufbau. Because of this it is quite puzzling how Friedman can think that taking structuralism into

\textsuperscript{10}This has of course been questioned by some philosophers, e. g. by Peter J. Lewis in [Lew01], who thinks that it contains a base rate fallacy. On the other hand, the no miracles argument has also been argued to contain a base rate fallacy. Therefore both of the assumptions at the base of Worrall’s argument for structural realism are dubious.

\textsuperscript{11}Russell’s causal theory of perception was different from current causal theories of perception. Current causal theories of perception are typically associated with direct realism about physical objects; however, Russell’s causal theory of perception was associated with representationalism. Russell held that the immediate objects of our perceptions were not the physical objects that ultimately caused them but sense-data, from which we inferred the existence of physical objects as their causes.
account should lead us to evaluate logical positivism more favourably.\footnote{Friedman does claim in [Fri87, pages 354,355] that Wittgenstein’s theory is not subject to the same objections as Carnap’s, which are of course the objections Newman presents. As we will see, in order to counter Newman’s argument one would have to abandon standard set theory, and Friedman claims that Wittgenstein does just that. However, as Friedman admits, one would have to pay a terrible price for this, namely the complete emasculation of classical mathematics. It is not clear to me if Friedman is willing to pay that price. However, I certainly do not think that it is a sufficient defence of structuralism that it would be saved by such drastic means; surely this is too high a price to pay for a more favourable evaluation of logical positivism.}

To make this problem clear we have to explain what we mean by a structure and by structural properties. The problem with many of the newer philosophers who call themselves structural realists, Worrall in [Wor89] included, is that they do not make clear what they mean by a structure. However, there is a very clear concept of structure that Russell and logical positivists used, and which is commonly used in mathematics and logic. If structuralism is based on this concept of structure, as Russell’s, Carnap’s and Schlick’s structuralist theories were, then Newman has shown as conclusively as any philosophical theory can be shown to be false that structuralism, as the claim that we can only have knowledge of structures, is utterly untenable.

Both Russell and Carnap begin their treatment of structures by considering relations. Russell says in [Rus27, page 249]:

Two relations $P$, $Q$ are said to be similar if there is a one-one relation between the terms of their fields, which is such that, whenever two terms have the relation $P$, their correlates have the relation $Q$, and vice versa.

Carnap says in [Car67, §11-12] that if two relations have the same arrow diagram, they are structurally equivalent or isomorphic. Russell adds in [Rus27, page 250]:

Two relations which are similar have the same ”structure” or ”relation-number”.

Newman on the other hand considers systems consisting of a set of objects $A$ together with a relation $R$ defined on $A$ i. e. holding between its members (i. e. $R \subseteq A \times A$) be given (see [New28, page 139]). Of course, in such cases $A$ is the field of $R$, so this is equivalent with the case considered by Russell and Carnap. Such systems (or just the relations contained in
them) are often (following M. I. G. Redhead) called concrete structures today (see e.g. [Vot, pages 42-43]), though neither Russell nor Newman yet used such terminology. The notion is very familiar for current logicians and mathematicians from model theory; a model in one common sense of the word consists of a concrete structure (often called a model structure) together with an interpretation function. Two such systems are said to have the same (abstract) structure, - the only kind of structure of which Carnap, Russell and Newman spoke, which Russell also called a relation-number - iff there exists a one-one correlation between them (i.e. if they are isomorphic; Russell used the word similar). In symbols if \( \langle A, R \rangle \) and \( \langle B, S \rangle \) are concrete structures, then they share the same abstract structure iff there is such a bijection \( f \) from \( A \) to \( B \) that for all \( x, y \in A \), \( R(x, y) \) iff \( S(f(x), f(y)) \).

Quine calls in many writings (e.g. in [Qui68, page 205] and in [Qui08a, page 404]) such functions \( f \) satisfying such a condition proxy functions.

Epistemic structuralism supposes that all we can know of systems consisting of unobservable i.e. theoretical entities is what abstract structures they have, and strong forms of it suppose that all we know of any entities, even observable entities, is what abstract structures they have.

---

13The notoriously ambiguous words "concrete" and "abstract" are used in a very special sense in drawing this distinction. According to most more usual senses of the words both kinds of structures would be abstract or both concrete. E.g. "concrete" often means being spatio-temporal, and if the word is used like this then an immanent realist like Armstrong would say that both kinds of structures are concrete, while a Platonist would say that both are abstract. It would be more perspicuous to speak of (purely) formal and (partly) material structures, but the use of the words "abstract" and "concrete" has become established so I will use these words.

14No abstract structures themselves can be sets according to standard ZF set theory - unlike Russell’s type theory, where they could be sets but would be divided into infinitely many structures of different types - though they can be proper classes in set theories that allow such entities and some of them can be sets in some non-standard set theories such as Quine’s NF and NFU. Similarly abstract structures cannot be properties or relations in higher-order logics whose axioms are similar to those of ZF. If we relax the latter restriction, then structures might be called higher-order properties or relations as they often are (e.g. in [Cha07, page 37]). It is important to note, however, that contrary to what many incautious exponents of structuralism often imply, all higher order properties and relations are not abstract structures; higher-order properties and relations can be non-structural i.e. material. It is only higher-order properties which are invariant under isomorphisms which are purely structural. The claim that we could only know higher-order properties of properties of theoretical objects would not be disproved by Newman’s argument, though I do not know if it could be motivated in any reasonable way.

15Epistemic structuralism has often, first by Grover Maxwell in [Max70], been associated with the Ramsey sentence approach to theoretical terms. In the Ramsey sentence approach the theoretical terms occurring in the conjunction of a theory’s axioms are replaced with free variables and then bound by a quantifier prefixed to this conjunction.
Newman pointed out (in [New28, pages 140, 144]) that according to standard set theory (and most non-standard set-theories) or standard higher-order logic (and most non-standard higher-order logics) any collection of entities can be organised so that they form a system belonging to any abstract structure if there are sufficiently many of these entities. In symbols if \( \langle A, R \rangle \) is any system and \( B \) any other set whatever with the same cardinality as \( A \), then there is such a relation \( S \) defined on \( B \) and such a system \( \langle B, S \rangle \) that \( \langle A, R \rangle \) and \( \langle B, S \rangle \) share the same abstract structure. Newman therefore showed that epistemic structuralism implies together with standard set theory that nothing can be known of theoretical entities except their number; any scientific theories that postulate the same number of physical objects would according to structuralism be equivalent (epistemically equivalent according to epistemic structuralism and ontologically equivalent according to ontic structuralism). Carnap had claimed in Aufbau (see [Car67, §15]) that there could be purely structural definite descriptions of individual objects, but Newman’s result shows that this is impossible (in any universe with more than one individual object, such as our universe obviously is). Nor can there be purely structural definite descriptions of any qualitative properties or relations or any classes other than structures either. Because of this, even if Friedman and Richardson were right that structuralism was more central to logical positivism than empiricism, this would not necessarily be any reason

However, it seems to me that the Ramsey sentence approach by no means implies structuralism (and hence need not share all its problems, though it may have similar problems of its own). It can be combined with stronger kinds of realism, as has been done e. g. by David Lewis in [Lew70]. This depends partly on how observational terms are understood; I will later divide them into weakly and strongly observational. If observational predicates occurring in the theory are weakly observational in the sense that while some of the entities belonging to its extension are observable, some are not, then the Ramsey sentence of the theory can quantify over unobservable entities. In the case of some theories, their Ramsey sentences can in this case carry ontological commitment not just to abstract structures but also to unobservable particulars and their intrinsic properties and qualitative relations. However, if the predicates are all strongly observational, then the Ramsey sentence approach collapses into the view that all we can know of unobservable entities is their number. Let me just give a fictional example. Let us suppose that a theory consists just of the statement \( T(a) \land (\forall x)(T(x) \rightarrow (R(x, b) \land Q(x))) \), where \( b \) is the name of a minimal observable particular and \( R(x, y) \) means that \( x \) is smaller than \( y \) and \( Q \) is an arbitrary (weakly) observational predicate denoting an intrinsic property. In this theory \( T \) and \( a \) are theoretical terms and \( Q \) and \( b \) and \( R \) are observational terms. The Ramsey sentence of the theory is \( (\exists X)(\exists x)(X(x) \land (\forall y)(X(y) \equiv (R(y, b) \land Q(y)))) \), and this sentence carries ontological commitment (just like the original theory) to an unobservable particular \( x \) and to an unobservable yet non-structural property \( X \), and actually tells us that the unobservable particular has an intrinsic property.
to re-evaluate logical positivism more positively.

In fact Newman’s results imply even worse news for the structural antirealism of the early logical positivists than for Russell or modern structural realists such as Worrall, for the early logical positivists were far more ambitious in their use of structuralism. Russell, Schlick (in the pre-positivistic phase of his philosophical development in \([\text{Sch}18]\)) and Worrall had thought that we can have only structural knowledge of unobservable entities, but had not limited our knowledge of observable or at least of observed entities in this way. They had allowed that we could know (in Russell’s case, of course, by acquaintance) non-structural properties of observable entities\(^{16}\). Newman’s results then implied that all the knowledge of theoretical entities their theory allowed was trivial, but they did not threaten their knowledge about observable entities. It allowed that they could know more about what they believed to be observable entities - sense-data in Russell’s case and macroscopic physical bodies in Worrall’s case - than how many there were. Carnap, however, had said in \(\text{Aufbau}\) (see \([\text{Car}67, \S16, \text{pages 28-30}]\)) that all scientific statements are structure statements, and held that this was necessary for the objectivity or intersubjectivity of science. As Friedman and Demopoulos notice in \([\text{DF}89, \text{page 635}]\), Carnap had attempted to go Russell one better by turning all constant terms into variables by means of purely structural descriptions. This implied that even statements about observable entities were structure statements, so we could not say even about observable entities anything except how many of them there were. Therefore if the objectivity of science is to be saved, it must be done in an entirely different way than Carnap tried to do. The same holds for the theory Schlick held in \([\text{Sch}26]\) (translated in \([\text{Sch}79a]\)); Schlick held there that only structure is communicable, but as we have seen, if that were all that was communicable, we could communicate nothing besides purely mathematical knowledge and how many concrete objects there are. Pointing to structuralism does not then give any reasons at all for evaluating the early Vienna Circle more positively than just attending to their empiricism, but rather far more negatively; in fact Carnap’s \(\text{Aufbau}\) becomes far more coherent if the structuralist elements are just removed from it.

\(^{16}\text{This was yet rather bad for Russell, since he thought that all physical entities were theoretical and only sense-data were observable. However, this conclusion would be less disastrous for someone who held physical entities to be observable.}\)
Newman’s results do not lead structural realism of the type that Russell, the pre-positivistic Schlick and Worrall defended to so absurd conclusions as they do the theory of Carnap in the *Aufbau* and Schlick (in his positivistic phase), but such structural realism is yet in truly serious trouble\(^\text{17}\). If the claims about the structure of theoretical entities are no more than trivial cardinality claims, then structural realism just collapses into a non-realistic theory such as phenomenalism or constructive empiricism, where only observable objects are known to exist. While Carnap may have welcomed such an outcome, Worrall would not have, since it leaves structural realism no advantage with respect to phenomenalism or constructive empiricism. The mere fact that there is a certain number of theoretical entities of which science can tell us nothing more cannot explain the empirical adequacy of science, but leaves it just as miraculous as anti-realism does.

We must recognize the limits of Newman’s argument; while it provides a conclusive refutation of structuralism in the original sense of the word, it of course does not refute every theory that has been called “structuralism”. Of course, since Worrall nowhere defined the notion of structure, he could evade the argument by claiming that his notion of structure is quite different from the common one used by Newman in his proof. However, in this case it becomes quite obscure what is meant by structural realism. The same holds of many other modern theories that are called structuralist\(^\text{18}\). However,

\(^{17}\)This does not of course imply that all causal theories of perception would be in serious trouble, for a causal theory of perception, even a representationalist one, need not be combined with structuralism as Russell combined them. In fact a representationalist causal theory of perception fits badly with structuralism, since it implies that the causal relation has unobservable entities as its members, and the causal relation is not plausibly understood as structural, so the representationalist causal theory of perception indicates that we can know at least one structural relation. Russell’s own theory could then be modified so that it was freed from structuralism. Actually there are many parts of Russell’s theory in [Rus27] that are likely to be incompatible with structuralism. E.g. the neutral monism he argues for in Chapter XXXVII of [Rus27], which claims that fundamental entities are neither mental nor physical, is incompatible with structuralism, since mentality and physicality are not purely structural properties, and therefore neither is the denial of their disjunction; however, clearly at least some fundamental entities are unobservable in Russell’s theory.

\(^{18}\)Ladyman, Ross and the others argue in [LwDSC07, page 128] that the Newman problem does not arise for ontic structural realism because it is not formulated in terms of quantificational first-order logic and in particular because they eschew an extensional understanding or relations without which the problem cannot in his view be formulated. However, Newman first proposed his argument against structuralism in connection with Russell’s theory which was not formulated in those terms. Russell’s structuralism was based on his and Whitehead’s theory in Principia Mathematica, which was a higher-order logic and was ultimately based on an intensional understanding of relations as proposi-
some philosophers who call themselves structural realists have formulated reasonably clear theories they hold to descend from structural realism which may not be vulnerable to Newman’s argument.

The crucial distinction here seems to be the one between concrete (or formal) and abstract (or material) structures. Any version of structuralism that admits that not only abstract structures, but also (at least some) concrete structures of unobservable entities - such as relations which are not invariant under all isomorphisms - exist and are objective and intersubjectively knowable, is immune to Newman’s argument.\(^{19}\) Such theories are still versions of selective realism and differ from full-fledged realism, as they deny intensional functions. This can be seen e.g. in [WR63, page xv], where they say that their system begins with atomic propositions which contain relations in intension and in [WR63, page 23] where they distinguish a propositional function from the class it determines. They do say in [WR63, page 22] that mathematics is only concerned with extensions, so they treat only extensional functions in [WR63]; however, this allows physics to be concerned with intensional functions (including intensional relations). However, Russell’s definition of structures by means of isomorphisms in [Rus27] makes all structures extensional. Therefore ontic structural realism cannot escape Newman’s argument just because it is not a first-order or extensional theory. However, it is not wholly clear whether this ontic structural realism is at all a structuralist theory in the sense in which Newman argued against structuralism, and if it is not, then Newman’s argument is of course powerless against it. Ladyman and his co-authors argue at length in chapter 3 of [LwDSC07] that there are no self-subsistent individuals but relations are ontologically more fundamental than their relata. This view, which seems to be most central for them does not yet imply structuralism in the sense of the word in which Carnap or Russell or Schlick were structuralists, for these ontologically fundamental relations do not need to be purely structural relations but can be qualitative. Ladyman and the others at least sometimes seem to mean with relational structures only concrete structures, not abstract structures, and if this is in the end correct then their theory is immune to Newman’s argument. However, at other times they do seem to make use of isomorphisms, and therefore of abstract structures, so parts of their theory seem to be vulnerable to Newman’s argument.\(^{19}\)E. g. it seems to me that Grover Maxwell in [Max70] included concrete structures in structures. He said in [Max70, page 188] that structural characteristics can be taken to be just those that are not intrinsic and can be described by means of logical terms and observation terms. He added that not all structural properties are also purely formal. An abstract structure, however, could be described with only logical terms and would be purely formal. Also Maxwell nowhere made use in this article of the notion of an isomorphism, which is essential to the definition of an abstract structure. Maxwell though himself to be explicating Russell’s concept of structure, but did not see that Russell’s concept was that of an abstract structure. To give a more recent example, Anjam Chakravartty has presented (e. g. in [Cha07, §2.3]) a theory he calls semi realism. He calls relations between first-order properties structures and admits that such relations can be qualitative. Newman’s argument against structuralism does not apply to semi realism (though it can be doubted whether semi realism is any less vulnerable to pessimistic induction than full-fledged scientific realism). However, since semi realism is explicitly presented as a metaphysical theory, and a realistic one, I do not need to resist it for it has nothing to do with logical positivists’ arguments against realism on the basis of their very different version of structuralism.
that we can have knowledge of unobservable individuals (and perhaps even of the qualitative intrinsic properties of unobservable entities). However, such theories are so far from what was originally called structuralism that it is rather misleading to apply the name to them. Such theories of course face many objections, but these are not as obviously insuperable as the Newman problem. I cannot here examine any further whether such theories are justified or not.

There have, however, been relatively recent structuralists, whose theory is clearly vulnerable to Newman’s argument. Quine’s structuralism in [Qui08a] seems to me to be such, as it makes use of proxy functions i.e. isomorphisms and therefore allows only knowledge of abstract structures. Furthermore, it seems to be vulnerable to Newman’s argument in the radical way in which Carnap’s and Schlick’s theories were rather than the less radical way in which Russell’s theory was, as it applies such isomorphisms not only to theoretical but also to observable entities (as I will discuss in more detail later).

We have then seen that as Newman showed, if any non-trivial knowledge is possible, it must be non-structural, and sciences must possess non-structural knowledge, so even if metaphysics requires non-structural knowledge, this is no reason to object to it.

3.1.4 Coherentist and Conventionalist Elements in Logical Positivism

The coherentist and conventionalist elements of logical positivism are closely associated with its structuralist elements. The reason for this is that the coherence of a theory depends on its structure rather than its content, and if we are to choose one of many equally coherent theories without any foundations this can only happen by means of conventions. These three theories together represent the anti-empiricist tendency in logical positivism.

The tension between the empiricist and non-empiricist elements of logical positivism was to some extent addressed already within the movement in a debate between Neurath and Schlick. Unfortunately, this debate was plagued by serious confusions. The epistemological question whether knowledge has foundations was regularly confused with semantical questions such the definition of truth. Schlick defended both foundationalism and the correspondence theory of truth already in his pre-positivistic period in an article
originally published in 1910 and translated into English in [Sch79b] and later as a full-fledged positivist in [Sch34] (translated in [Sch59]). Neurath on the other hand apparently defended both a coherence theory of justification and a coherence theory of truth (e.g. in [Neu32] - translated in [Neu59] - and in [Neu34]), though there is some dispute about the latter. Of course, these views are not necessarily connected in any obvious way; it can be argued that one can also consistently be a correspondence theorist about truth and a coherentist about justification.

However, this confusion may not matter much, as coherentism is not a plausible theory of either truth or justification; the claim that any arbitrary coherent fairy tale is justified is not much more sensible than the claim that any arbitrary coherent fairy tale is true. While the correspondence theory of truth does not imply that we should be capable of comparing truthbearers with reality, foundationalism does, and Schlick argues in my view quite plausibly that we can in fact compare them with it. Another confusion is more fatal.

Schlick and Neurath both presupposed that foundationalism led to the view that absolutely certain knowledge was possible; they did not see that a weak, fallibilistic foundationalism was possible. See e.g. [Haa09, page 54] for a distinction between weak and strong foundationalism, where strong foundationalism requires conclusive aka decisive justification, while weak foundationalism requires only defeasible, prima facie justification. Bertrand Russell, who had began as a strong foundationalist, retreated in his later philosophy to weak foundationalism (if indeed his theory remained foundationalist at all and did not anticipate foundherentism) and defended it in arguing against Neurath, saying in [Rus40, page 150] that

a non-inferential belief need not be either certain or indubitable.

Schlick begins [Sch34, page 79] (translated in [Sch59, page 209]) by arguing that all the important attempts at establishing a theory of knowledge grow out of the problem concerning the certainty of human knowledge. However, the real problem Schlick eventually has with Neurath’s view in [Sch34, page 86] (translated in [Sch59, page 215]) is not that it does not guarantee certainty but that it makes the choice of theories entirely arbitrary, so that one would have to consider fairy stories to be as true as a historical report. A weak foundationalism suffices to make the choice of theories non-arbitrary.
and non-conventional although it does not offer incorrigibility. On the other hand Neurath thought falsely that rejecting the arguably implausible position that human knowledge could be incorrigible required the total rejection of foundationalism. There is then a third option besides the views of Schlick and Neurath.

Extreme conventionalism has also serious problems, just like the coherentism with which it is often associated. Friedman sees in it as a great achievement of Carnap that he extends Poincaré’s conventionalism to logic itself. However, that very extension has been revealed to be an especially problematic variation of conventionalism. While without that extension conventionalism is just a corollary of coherentism, this extension makes it a stronger and hence more absurd claim that coherentism. Quine argued in that the view of many logical positivists that even logical truths were conventionally true leads to a vicious infinite regress (even though Quine was in that article willing to entertain the rather extreme view that set theory was conventional). The basic argument against logical conventionalism is quite simple. As logical truths are infinite in number, they cannot be stipulated one by one, but must be deduced from some general conventions by means of some logical rules. However, in so doing we are presupposing the very logical rules or truths we are trying to establish by means of conventions.

There is a still greater problem with logical conventionalism. If we assume that the transformation rules (i.e. inference rules) of a language can be chosen quite arbitrarily, as Carnap does in formulating his Principle of Tolerance according to in , we can (as Arthur Prior showed

\[\text{These confusions persist in commentators dealing with logical positivism. Coffa still thought in that fallibilism was equivalent with the rejection of foundationalism and (in that accepting Tarski’s theory of truth would in some way count against a coherence theory of justification. Similarly, Misak says in that the foundationalist is after a foundation of rock, not of shifting sand. This, however, only holds of strong foundationalists, not of weak foundationalists. Of course, everyone would prefer a foundation of rock. However, it may be that human knowledge just cannot have such a foundation, and a foundation of shifting sand is better than no foundation at all.}

\[\text{This argument of Quine’s is quite independent of the arguments in that Friedman saw as ending logical positivism; in Quine sees analytical statements that are not logical as unproblematic, while in he turns his attack on them. Indeed, Quine’s views in the two articles may not fit very well together; the infinite regress argument against logical conventionalism is quite similar to an infinite regress argument other philosophers (e.g. Laurence BonJour in have used against the extreme empiricism Quine argued for in .}

\[88\]
introduce a connective tonk, such that by using the transformation rules of that connective we can deduce any sentence from any sentence whatever. We can then admit any sentence whatever, no matter how much it is in contradiction with experience. Here logical positivism collapses into extreme relativism.

I will later return to the discussion of the conventionalist element in logical positivism when discussing Neurath’s version of verificationism; I will show that in Neurath’s treatment verificationism loses all its force and collapses into conventionalism and hence into relativism.

3.1.5 Syntacticist attacks on Metaphysics

Many of the earliest logical positivists thought that philosophical problems were linguistic problems and furthermore problems which concerned the syntax of the language of science, not its semantics. This syntacticism can be viewed as special case of structuralism, an especially narrow case, as it allowed only syntactic structure, not semantic, to be relevant for philosophy. It played the same role in their philosophy as psychologism played in earlier philosophies, in British empiricism and German idealism: to avoid traditional metaphysical problems suggested by a realistically interpreted logic and mathematics. A special form of this syntacticism restricted to the philosophy of mathematics and logic was formalism, famously defended by David Hilbert, against which Frege already strongly protested, but which nevertheless was embraced by many analytical philosophers though they claimed to follow Frege. David Hilbert was associated with the Berlin Circle of logical positivists, so they were influenced by Hilbert’s ideas, which they generalized. Wittgenstein’s famous claim that philosophical problems were grammatical problems was probably an obscure statement of a generalized and extreme form of this view. A generalized form of this point of view was famously presented in a more exact form by Rudolf Carnap in [Car37].

Carnap claimed at the very foreword of this work [Car37, page xii] that philosophy is to be replaced by the logic of science and that the logic of science is nothing other than the logical syntax of the language of science. Thus logic need according to him have nothing to do with semantics and hence need not concern itself with ontological issues either, and since philosophy is nothing more than logic it need have nothing to do with ontology either\textsuperscript{22}.

\textsuperscript{22}Carnap distinguished in [Car37, §21, page 6-7] between pure and descriptive syntax,
Carnap thought that this syntactic character was what distinguished philosophy from the natural sciences, but as we have seen Neurath’s coherentism implied that even natural sciences could only make syntactic claims of what statements followed from what theories by inference rules.

In applying syntacticism to traditional philosophical problems in \[\text{Car37, §73, page 281}\], Carnap tried to show that many questions - those traditional philosophical questions which were not really meaningless - which seemed to concern non-linguistic objects were really only pseudo-object-questions. He argued that these questions appear to refer to objects because of a misleading formulation but really only refer to linguistic expressions. The sentences which are possible answers to those questions were then according to him (see \[\text{Car37, §74, pages 284-287}\]) also pseudo-object sentences, which seemed to talk about non-linguistic objects but really talked about language. They must then be translated from the material mode of speech into the formal mode of speech, i.e. into syntactic questions. This doctrine was later changed by Quine to the more defensible if still controversial form of a doctrine of semantic ascent, where it was still held that philosophical sentences often ought to be translated into linguistic questions to clarify them, but it was no longer held that such translation resulted in syntactic sentences, but instead semantic ones.

Of course if Carnap’s view of the exclusive significance of syntax were true, it would have enabled analytical philosophers to study language wherever pure syntax is wholly analytic while descriptive syntax is concerned with the syntactical properties and relations of empirically given descriptions. Carnap’s thesis that philosophy is syntactical acquires even minimal plausibility only because he in practice neglected this distinction of his own. It is indeed plausible (though already controversial) that pure syntax would be a part of logic, though scarcely that it would be all of logic. However, descriptive syntax cannot plausibly be regarded as belonging to logic at all. Carnap said later in the concluding chapter of his book \[\text{Car37, §86, page 332}\]:

Our thesis that the logic of science is syntax must therefore not be misunderstood to mean that the task of the logic of science could be carried out independently of empirical science and without regard to its empirical results. The syntactical description of a system which is already given is indeed a purely mathematical task. But the language of science is not given to us in a syntactically established form; whoever desires to investigate it must accordingly take into consideration the language which is used in practice in the special sciences, and only lay down rules on the basis of this.

Here Carnap practically admitted that philosophy, even if limited to the philosophy of science, must include not only pure but also descriptive syntax; if philosophers were concerned only with pure syntax they would have no reason to take into consideration the language used in the special sciences. However, since such descriptive syntax is not a part of logic in a proper sense, neither is the philosophy of science.
out considering its relationship to the rest of the world. However, things changed when Gödel proved his incompleteness theorems in [Göd31], which showed the untenability of formalist views in mathematics and even more when Tarski developed his theory of truth and his theory of logical consequence, which made logical consequence a matter of semantics instead of syntax. The falsity of the syntacticist view was then temporarily widely recognized. Even Carnap altered his view because of Alfred Tarski’s theory of truth in 1942 in [Car48]. Carnap said now [Car48, §39, page 246] that

The field of theoretical philosophy is no longer restricted to syntax but is regarded as comprehending the whole analysis of language, including syntax and semantics and perhaps also pragmatics.

Carnap said also that the task of philosophy is semiotical analysis, where semiotics contains besides syntax also semantics and pragmatics.

Carnap was unusually honest for a philosopher in admitting that he had really changed his views; nevertheless even he did not quite admit the full extent of the change. Carnap said that many of the earlier analyses are now seen to be incomplete although correct. While this holds of many details, which remain correct and even of permanent value, the general thesis of the earlier work was surely completely abandoned; not merely modified but replaced with its polar opposite. In fact Carnap’s new view would have compelled him to abandon positivism itself, to reject all the theses which were distinctive of logical positivism, if he had been consistent.

Most analytic philosophers now allowed that semantics is possible and is an indispensable part of the logical study of language. Since such Tarskian semantics requires a consideration of the general relations between the structures of the language and the world, and thus a consideration of the general structure of the world, i. e. of formal ontology, one would have thought that this would have led to the restoration of ontology. However, many philosophers still continue to assume uncritically that the study of language and logic is somehow possible apart from the study of the world as a whole, apart from ontology.

Carnap himself never abandoned his opposition to metaphysics, even though all the reasons he had originally given for it were abandoned, which shows that the opposition to metaphysics was an irrational part of Carnap’s
otherwise rational philosophy. He still tried in [Car48, page vii] to reassure his fellow empiricists, who fear that a discussion of propositions as distinguished from sentences and of truth as distinguished from confirmation will open the back door for speculative metaphysics, that he would not like to help in reviving it. He said that the problems of philosophy concern not the ultimate nature of being but the semiotical analysis of the language of science.

However, all such reassurances were empty and unjustified. Carnap had earlier distinguished metaphysical questions which he regarded as meaningless from the meaningful philosophical problems by the very feature that meaningful philosophical problems were syntactical problems. Having rejected syntacticism, Carnap no longer had any way of distinguishing metaphysical questions from what he thought were genuine philosophical problems in any principled way such as he once thought he had. He could scarcely go back to verificationism since the new semantical statements he now regarded as meaningful were not always verifiable. The new semantical theory contained such claims as that predicates designate properties. This claim clearly implies such a statement as that there are properties. This is surely a claim about the ultimate nature of being.

Of Carnap’s fellow positivists, Otto Neurath famously attacked (as Carnap tells in [Car63a]) Tarski’s theory of truth and Carnap’s defense of it (as Bergmann in his positivistic period also did in [Ber44]) and claimed that the semantical conception of truth was incompatible with the anti-metaphysical view which was characteristic of positivism. If what I have argued so far is correct, then Neurath was of course quite correct that there was such an incompatibility. Carnap certainly did not succeed in showing that the Tarskian theory would be compatible with positivism. However, an incompatibility between two theories can always be used against either of the theories by those who believe in the other. Thus also in this case while the incompatibility can be used against formal semantics as Neurath did, it can also be equally well be used against the anti-metaphysical view. Since Tarski’s semantical theory has in the view of many philosophers proven itself to be quite fruitful, while all the arguments the logical positivists had for the impossibility of metaphysics have been shown to be based on incoherent or dubious premises, many philosophers today would rather reject positivism than formal semantics.
Not only does the move to the semiotic theory deprive Carnap of any reason to think metaphysics impossible; it actually drives him to investigations which can themselves with good reason be held to be parts of metaphysics, though he arbitrarily refused to count them as such. Carnap distinguishes in [Car48, §17, page 88-95] absolute concepts from semantical concepts. According to him for every semantical concept $M_s$ attributed to $n$ expressions there is a corresponding concept $M_a$ such that whenever $M_s$ holds for $n$ expressions $M_a$ holds for the designata of these expressions\textsuperscript{23}. These absolute concepts are not semantical concepts although they are closely related to semantical concepts. They belong to the non-semiotical part of the metalanguage. Such absolute concepts are thus clearly not linguistic concepts at all, as they can be used without referring to any languages at all, even though they are useful in linguistics. However, many absolute concepts, especially those corresponding to what Carnap called pure semantics, and especially those which Carnap uses in his exposition, are so general that they cannot be concepts specific to any special science either. They are thus clearly purely metaphysical concepts. It follows that statements in which only such absolute concepts occur are purely metaphysical statements and theories consisting entirely of such statements consisting solely of absolute concepts are purely metaphysical theories. While the rest of claims in Carnap’s semantics are just implicitly metaphysical, those made using such general absolute concepts are explicitly metaphysical. Carnap defines a correspondence between semantical and absolute concepts; this is in fact a correspondence between semantical and metaphysical concepts and establishes also a correspondence between the statements of semantics and metaphysics.

A metaphysics based only on the absolute concepts Carnap himself explicitly mentions would of course be far too thin and poor to count as a serious metaphysical theory; however, the correspondence defined by Carnap can be applied to any semantical concepts to produce new absolute

\textsuperscript{23}We can formalize this correspondence by the following formula: $C(M_s, M_a) \equiv (\forall M_s)(\exists M_a)(\forall x_1)\ldots(\forall x_n)(M_s(x_1, \ldots, x_n) \rightarrow M_a(D(x_1), \ldots, D(x_n)))$. Oddly enough Carnap later formulates the correspondence in a different way which inverts the direction of implication: Carnap says that a term used for a radical semantical property of expressions will be applied in an absolute way (i.e. without reference to a language system) to an entity $u$ if and only if every expression $U_i$ which designates $u$ in any semantical system $S$ has that semantical property in $S$. It seems to me that Carnap’s earlier definition of correspondence is definitely better.
concepts. If the semantical concept in question just is sufficiently general (like most semantical concepts that are naturally introduced in the context of semantical research), the corresponding absolute concept will be a metaphysical concept. In this way a quite rich metaphysics can be constructed. Since Carnap works mainly in type-theoretical languages, he does not take the concepts of attribute and proposition themselves to be absolute concepts. However, in a stronger language, such as Carnap called in [Car48, §12, page 51] a language with transfinite levels or a language system without distinctions of types or levels, these concepts would also be absolute concepts.

No matter how unwillingly, Carnap had flung the back door wide open and questions of metaphysics must inevitably enter through it. However, happily they may not turn out to be as unwelcome visitants as he thought, but the new exact methods of semantics may help to make them every bit as clear as the questions of natural science seemed to him.

Similarly to Carnap, other philosophers have also tried to go around the results of Gödel and Tarski in some way, usually a quite obscure one. Syntacticism has again become popular and the lessons of Tarski and Gödel have been forgotten. This attitude has been partly rationalized by developing Tarskian semantics in a less realistic direction. I will argue that modern model-theoretical semantics represents philosophically a retrograde step from the insights of Tarski and Gödel back towards the confusion of for-

\[\text{\scriptsize 24}\]

Michael Friedman, who as we have seen emphasized the structuralist elements in the thought of Carnap and other logical positivists, sees the syntacticism of [Car37] as the further development of such structuralism, as it too is committed to a formalistic and holistic theory. Friedman concludes in [Fri87, page 540] that

the anti-metaphysical dream of Vienna finally stands of falls, therefore, not with phenomenalism, but rather with the remarkable program of Logical Syntax itself.

Since Carnap himself abandoned the program of Logical syntax in [Car48] as a result of Tarski's results, we can surely say that in that case the dream definitely falls, with a loud resounding thump. Many philosophers, however, for whom it was a nightmare rather than a dream, will be glad for this. It is weird why Friedman and other proponents of the new interpretation of Carnap and other positivists do not often discuss [Car48] at any length; it is usually only briefly mentioned and then ignored. Perhaps they just ignore this work because it does not fit their interpretation; most of the ideas presented in it are not inspired by either empiricism or Neo-Kantianism and are in fact quite incompatible with either. Nevertheless these ideas also may have a longer history behind them. They seem to represent a third and wholly distinct tradition, perhaps one ultimately inspired by Austrian and Czech metaphysical realists like Bolzano, as e. g. Jan Wolenski has sometimes suggested.
malism. While Tarski in his original theory of truth studied the absolute denotation and truth of expressions, modern model-theoretical semantics always, even when there is no reason, relativizes their interpretation to models and thus sidesteps the most significant ontological questions, questions which the study of absolute denotation inevitably brings up. Of course, the existence of models already has itself ontological consequences, so even model-theoretical semantics does not completely avoid dealing with metaphysical questions; however, verbal quibbling is used to obscure even these remaining questions. Models consist of sets, and it is claimed that sets are not a part of the world in an ontologically significant sense, so that assuming them to exist does not involve making any ontological commitments. When dealing with the question of ontological commitment, I will argue that this is an untenable view. If there are sets, then they are part of the world in the wide sense in which the world is the object of ontology (if indeed there is any world), for in this sense the world is just the totality of everything there is. In any case, if we return to Tarskian semantics and seek for the absolute interpretation of expressions, then many expressions intuitively have as their semantic values entities which are not sets and in any case belong to the world.

3.2 Logical Positivists’ Arguments against Metaphysics and Their Problems

Logical positivists usually did not familiarize themselves with the metaphysical traditions that they criticized before criticizing them, so they did not interpret the metaphysical methods or claims that they attacked correctly, which already renders most of their criticisms of these metaphysical claims worthless. Positivists such as Carnap in [Car32b, page 229] often used Martin Heidegger, specifically Heidegger’s [Hei43], as a representative of metaphysical nonsense. However, Heidegger’s philosophy was by no means a representative example of traditional metaphysics, but there are reasons to think that it was rather a protest against it almost as much as the philosophy of logical positivists. Heidegger talked (e. g. in [Hei43, page 9]) about the overcoming of metaphysics (Überwindung der metaphysik) just as Carnap did in the very title of [Car32b], using the very same German word ”überwindung”. Heidegger may indeed have used the word in a slightly dif-
ferent sense and unlike Carnap he does not seem to have completely denied the legitimacy of metaphysics. However, it is clear that Heidegger was also highly critical with respect to metaphysics. Heidegger blamed metaphysics for thinking only generally about beings and forgetting Being itself, the Truth of Being and its essence (Wesen) which he said is distinct from metaphysics itself. Heidegger’s thinking tries to go back to the ground (Grund) of metaphysics. It is then relatively clear that Heidegger was not himself trying do metaphysics in any traditional sense in [Hei43], though it is far from clear what he was trying to do instead. In this situation, even if Heidegger’s sentences are nonsensical\textsuperscript{25}, there is no reason to suppose that their nonsensicality resulted from their being results of a metaphysical investigation pursued by a specific metaphysical methods, since Heidegger was not trying to do metaphysics when enunciating them.\textsuperscript{26}

Because of this susceptibility of logical positivists to the charge of misrepresenting their target many metaphysicians would claim that even if an acceptable version of the verification principle would be found, many metaphysical claims (once their essential meaning has been freed by semantical analysis from unnecessary traditional encumbrances) could survive a verificationist test, especially once the verification principle has been weakened enough by the lengthy process of constant criticism and reformulation.

A historically influential example of logical positivists’ attack on metaphysics is the English logical positivist Alfred Jules Ayer’s purported demonstration of the impossibility of metaphysics in [Aye34], a demonstration also incorporated with revisions into Ayer’s magnum opus, the English manifesto for logical positivism [Aye36a]. Ayer says that the views he expresses are not original, but derived from such earlier positivists as Wittgenstein, Schlick (in [Sch32b], translated as [Sch49b]) and Carnap (in [Car32b]). I will

\textsuperscript{25}It is indeed quite plausible that many of Heidgger’s formulations, such as the classic "Das Nichtsichtet" or "The Nothing nihilates." are meaningless. However, you do not need verificationism to argue that they are meaningless. Even someone who thinks that any grammatical sentences whose basic terms have ostensive definitions are meaningful can agree that many of Heidegger’s sentences are meaningless, as they are not even grammatical, at least according to many syntactical theories.

\textsuperscript{26}Recently many existentialists and deconstructionists have tried to develop an anti-metaphysical philosophy on the basis of Heidegger’s ideas. Of course, it can be argued that behind the front of such a purportedly anti-metaphysical philosophy you can often find an implicit metaphysics which is usually bad just because it is only implicit, and with many deconstructionists it is usually an obscure and outdated metaphysics, typically a Hegelian metaphysics or some other metaphysics inspired by German Idealism.
mainly concentrate on Ayer’s version of the attack, but also refer to these earlier continental sources of Ayer’s argument.

Ayer defined a metaphysical enquiry as

an enquiry into the nature of a reality underlying or transcending the phenomena which the special sciences are content to study

and said in [Aye34, page 339] that

it is the aim of metaphysics to describe a reality lying behind experience.

The great problem with these definitions is that Ayer did not even refer to any actual metaphysician who would have defined metaphysics in this way, much less show that all philosophers who are generally accepted as metaphysicians would have accepted such a definition. In fact all of Ayer’s scanty examples of metaphysical claims seem to derive (if only by hearsay) from the tradition of post-Kantian German idealism. Ayer did not quote any actual metaphysician in [Aye34], and in [Aye36a] he in his actual argument for the meaningfulness of metaphysics only cited Bradley, an extreme representative of English Hegelianism, of a kind of monistic idealism often called Absolute Idealism. Earlier analytical philosophers had of course also criticized Bradley without attacking metaphysics as a whole, so Bradley was an easy target. At least Bradley is clearly a genuine metaphysician, unlike Heidegger, whom Carnap took as his example, though Bradley is scarcely a typical metaphysician either. Ayer even claimed (in [Aye36a, page 38]) that George Berkeley was not a metaphysician, which makes his conception of metaphysics appear very different indeed from any common one. There then appears to be a danger that Ayer is just attacking straw men.

However, a little bit of historical research shows that some metaphysicians would have accepted that metaphysics is such an enquiry as Ayer describes. We must first of all look at how Bradley, whom Ayer cites as an example, defines metaphysics. Bradley follows Aristotle’s bad example by suggesting three different and by no means obviously equivalent definitions or characterizations. These characterizations, however, are not the same as Aristotle’s characterizations, but introduce among them the kind of definition referred to be Ayer. Bradley says in [Bra93, page 1]:

We may agree, perhaps, to understand by metaphysics an attempt to know reality as against mere phenomena, or the study
of first principles or ultimate truths, or again the effort to comprehend the universe, not simply piecemeal or by fragments, but somehow as a whole.

The first of these characterizations agrees rather well with the definition given by Ayer, but only the first. Therefore even if Ayer had succeeded in showing that metaphysics in the sense of the first definition was impossible, this would not yet have amounted even to showing that all that Bradley was trying to do was impossible.\textsuperscript{27} Even more importantly, even if the definition partially fits Bradley’s theory, there is no reason to think that the definition would fit in any way all theories that are commonly regarded as metaphysical. While Ayer was not quite attacking a straw man, he came close, for he attacked the weakest possible representative of metaphysics. Earlier analytical philosophers had already attacked resoundingly the metaphysical systems of Bradley and other absolute idealists without attacking all metaphysics.

If we instead use the traditional definition of metaphysics as the science of being qua being, which I have argued to be the most fundamental of the original Aristotelian definitions of metaphysics, Ayer’s demonstration does not work, for metaphysics in this sense also deals with experienced

\textsuperscript{27}A similar characterization can be found in earlier thinkers. Pierre Duhem’s conception of metaphysics and its relation to natural science in [Duh91, page 10] would seem to fit well with Ayer’s definition. Duhem says:

Now these two questions - Does there exist a material reality distinct from sensible appearances? and What is the nature of this reality? - do not have their source in experimental method, which is acquainted only with sensible appearances and can discover nothing beyond them. The resolution of these questions transcends the methods used by physics; it is the object of metaphysics.

However, Duhem did not actually define metaphysics in such a way as Ayer did. In any case, Duhem is scarcely a hard-core metaphysician, but has rather often been counted among positivists; though he thought that metaphysical questions were meaningful he was sceptical about the possibility of answering them. Therefore that Duhem accepted this definition does nothing to show that it would characteristic of metaphysics in general. Duhem did not take unverifiable statements to be meaningless, but thought yet that scientific statements had to be verifiable. This is similar to Popper’s later view of falsifiability; Popper took unfalsifiable statements to be meaningful but not scientific, just as Duhem thought unverifiable ones to be such. While these views would seem to leave some room for metaphysics, they yet erect an implausibly deep a gulf between natural science and metaphysics; a naturalistic metaphysician who thinks metaphysics is continuous with sciences could not accept such a gulf. Nor is there any reason to accept such a gulf since is easy to argue (as I do in this section) that on the one hand natural sciences must already often over-leap the bounds of verifiability and falsifiability while on the other hand some metaphysical statements are conclusively verifiable or falsifiable if any statements are.
objects (or objects that are constituents of experience), since they also exist. Since metaphysics in this sense must provide a category under which every entity falls, it must also provide a category for every experienced entity, with the possible exception of those that are purely hallucinated and do not really exist. Therefore, unless we assume absurdly that all experience is hallucinatory, metaphysics as the science of being qua being deals also with experienced objects. Even more, as I will show soon, Ayer’s theory implies that only objects that can be contents of experience i.e. phenomena exist (in the fundamental sense of the word ”exist”), so a theory that dealt with all phenomena would deal with all that there is, so Ayer’s total theory does not imply that any question of metaphysics as the science of being as being would be meaningless or even unanswerable. Also if we use the definition of metaphysics as the science of fundamental causes and principles i.e. metaphysics as aetiology then, as I will show, not only is even the strong form of verificationism consistent with the possibility of metaphysics in this sense, but Ayer’s own writings contained a metaphysical theory in this sense.

In fact the only one of Aristotle’s definitions that might seem to speak about the transcendent is the definition of metaphysics as the science which deals with that which exists separately. A separate (or separable) entity can indeed by said to be transcendent to those entities from which it is separate (or separable) (and indeed Plato’s ideas, which Aristotle thought to be separate or separable, have often been called transcendent by those who accepted Aristotle’s interpretation of Plato). However, even this definition is probably not equivalent or coextensive with Bradley’s and Ayer’s definition, at least relative to Ayer’s theory of experience, though the matter is not wholly clear (and Bradley’s and Ayer’s definition might have evolved historically out of this definition). Aristotle meant with separateness separateness (or separability) from matter, not separateness (or separability)
from experience, so the objects of metaphysics were by him supposed to be
by definition transcendent to matter, not transcendent to experience. The
sense contents that Ayer takes as fundamental can in a sense be said to
immaterial (though hardly divine, as Aristotle supposed separate entities to
be!), and as we will also see Ayer thinks that they could exist even if ma-
terial objects did not. Ayer’s theory might then be said to be metaphysical
even in this sense, to concern entities separable from matter.

Summing our results up, we have shown that it is very likely that none
of Aristotle’s definitions of metaphysics, which first marked out that dis-
cipline, agrees with Ayer’s definition. Therefore even if the verificationist
principle showed that such an enquiry as Ayer defined would always result in
meaningless statements, this would not suffice to show that all metaphysical
statements in the ordinary sense of the word were meaningless (though in
fact, as I will show, only a strong version of the verificationist principle, such
as Ayer himself eventually rejected, would have shown even this much).

Ayer completely ignored the long Aristotelian metaphysical tradition,
continental rationalism etc. One might try to defend Ayer by saying that
he thought them to be irrelevant to the philosophy of his day and attacked
the kind of metaphysics that was then relevant. However, this kind of de-
fence would not excuse the fact that Ayer ignored the way earlier analytical
philosophers had defined metaphysics. As we have seen, Bertrand Russell
defined metaphysics at the very beginning of [Rus18b, page 1] as the at-
tempt to conceive the world as a whole by means of thought, not as an
enquiry into the nature of a reality underlying or transcending the phenom-
ena or experience. Therefore just like Kant’s criticism Ayer’s attack does
not even threaten to touch the work of those many modern metaphysicians
who are Neo-Aristotelians or of metaphysical naturalists or of analytical
metaphysicians. Ayer’s definition really fits well only metaphysics derived
from post-Kantian German Idealism, though it also fits to some extent such
earlier traditions as Eleatic and Platonistic metaphysics (and Schlick does
indeed mention Plato and the Eleatics in [Sch32b]), which also viewed the

would probably not be. When Aristotelians called material bodies sensibles, this was only
because some of them were sensible, and did not imply that they would have denied the
existence of unobservable bodies. It can be argued that some bodies Aristotelians thought
existed such as the ether and crystalline spheres, were according to them unobservable,
though this is not wholly clear. At least such philosophers of science as Laudan and Wor-
rall (see [Wor89, pages 112,113]) have assumed that ether would have been unobservable,
though they have not given any historical evidence for such a view.

100
sensible world as mere appearance; however, these philosophers could not have taken this as a definition of metaphysics, as they did not (for all we know) even have a notion of metaphysics as a separate discipline, since they probably did not even have the idea of any separate scholarly disciplines at all; for them there was yet only philosophy as an undifferentiated whole. In presupposing that special sciences are content to study phenomena Ayer already smuggles in idealist metaphysical theories to his purportedly non-metaphysical conception, since that view (if the word "phenomena" is used with any exactness) derives from Kant and the German idealists (which is of course no accident, since the German logical positivists Ayer follows were originally Neo-Kantian). A more realistic interpretation of natural science, such as would be connected with a realistic metaphysics, would say that the special sciences already attempt to study things in themselves.

Curiously, one of Ayer's sources, Schlick, had earlier held this very view, and indeed in a stronger form. In [Sch26, page 152], translated in [Sch79a, page 105], Schlick actually says (in apparent complete contradiction to Ayer's view) that if we define metaphysics as the science of the transcendent, it is not merely possible, but the easiest thing in the world, since every proposition that we utter at all has a meaning that extends beyond the immediately given or experienced. Schlick's radical claim that we can only know the transcendent does not seem any more plausible to me than the view that we cannot know the transcendent at all, and I suspect it has equally disastrous consequences, leading to a strong structuralism which is not compatible with Schlick's own empiricism and which Newman has shown to be untenable. However, what is more pertinent now is that Schlick saw in 1926 that this definition of metaphysics used by Ayer was not a good definition, since metaphysics was supposed to be a special science, different from other sciences.

Schlick went on to claim that metaphysics has been used only for intuitive knowledge of the transcendent. This shows that Schlick followed Kant in erroneously assuming that metaphysics was supposed to be separated

---

29The contradiction between Schlick and Ayer is perhaps not as strong as it seems at first sight, since Schlick in [Sch26, page 151] (see for a translation [Sch79a, page 104]) calls non-experienced objects transcendent, regardless of whether they are viewed as logical constructions or whether they are credited with independent reality. This seems an unusually weak definition of transcendence, and Ayer might have been employing a stronger (and more natural) conception of transcendence, according to which logical constructions out of experienced entities are not transcendent.
from other sciences by a special mode of knowledge (intuition), not only by its object. Schlick can easily show that such knowledge is impossible, if we accept a conception of experience such as his, according to which something of which we have intuitive knowledge must be a part of the very stream of consciousness. However, if the act-object analysis of consciousness and experience which Schlick rejects were instead adopted (as non-positivistic analytical philosophers had adopted it), and it were thought that even in the case of intuitions there is a distinction between the act and its object (as these earlier analytical philosophers often thought), then even intuitive cognition of the transcendent might also be called “the easiest thing in the world”. For example, G. E. Moore said in [Moo22c, page 27] that merely to have a sensation is already to know something which is not a part of experience. Sensations in Moore’s sense can be taken to be intuitions, since Moore contrasts them in [Moo22c, page 7] with thinking, which is usually contrasted with intuitions. However, something that is not part of experience is transcendent to experience in the sense in which early logical positivists like Ayer used the word. The view that we can have intuitive knowledge of the transcendent had then been held not only by wild speculative metaphysicians such as Schopenhauer and Bergson but also by sober pre-positivistic analytical philosophers such as Moore. Many modern post-positivistic analytical philosophers have also held this view, including metaphysicians like Armstrong, who is a direct realist about physical objects,$^{30}$ holding that they are objects of immediate perception (which has often been taken as the definition of intuition). Many other philosophers, such as Edward Pols in [Pol92], have also held such a view. If Moore was correct or if such philosophers as Armstrong and Pols are correct,$^{31}$ then intuitive knowledge

$^{30}$ Armstrong says in [Arm61, page xi] that Representationalism and Phenomenalism hold that the immediate object of awareness is a sense-impression or sense-datum and that such an object, it is usually assumed, cannot exist independently of the awareness of it. If this were correct, then representationalists and phenomenalists would not accept the possibility of an intuitive cognition of the transcendent. However, while Ayer did indeed assume what Armstrong says of his sense-contents (and therefore is a phenomenalist of the kind Armstrong refers to), Moore, who is one of the original discoverers or inventors of the notion of sense-datum dis not. Though Moore was (at least at many phases of his philosophical development) a representationalist in a sense, as he thought that the knowledge of physical objects was indirect, he yet thought that sense-data could exist independently of the awareness of them.

$^{31}$ A big part of why the early Schlick thought that intuitive knowledge of the transcendent is impossible was that he thought it was not needed. This in its turn is because he thought that our knowledge of the transcendent could be purely structural. However, as
of the transcendent is indeed possible but is something that every discipline contains, so Schlick’s definition does not work to distinguish any separate discipline either.

In any case, Schlick’s definition of metaphysics seems to me just as arbitrary as Ayer’s definition; it has no historical justification. It does not for example correspond to any of Aristotle’s definitions, which say nothing whatever of intuition. Aristotle probably indeed held that his primary principles and causes are transcendent to experience, and it is highly plausible even to most non-Aristotelian metaphysicians that metaphysics as aetiology in fact deals with the transcendent. However, Aristotle’s definitions of the discipline allow also for metaphysical theories which take the fundamental entities to be immanent, and indeed various kinds of subjective idealism and neutral monism are such immanentist metaphysical theories. Even more importantly, many entities which are not primary causes and principles (e. g. molecules) are also naturally analysed as transcendent, and are not dealt with by metaphysics but by the special sciences. It is unlikely that Aristotle or original Aristotelians thought that humans could have intuitive knowledge of the Intelligences or prime matter he posited, though Neo-platonists and later Aristotelians influenced by them thought that possible. Schlick referred to Schopenhauer and Bergson, and it is indeed likely that they had held metaphysics to be intuitive knowledge of the transcendent; however, it is not sure that they would have accepted Schlick’s conceptions of transcendence and intuition, so it is not clear that what they tried to do was impossible. In any case they are not typical of all metaphysicians any more than Bradley, so even if intuitive knowledge of the transcendent were impossible, it does not follow that all metaphysics in the more common senses of the word would be impossible, but only that Schopenhauer’s and Bergson’s specific metaphysical theories were incorrect. Schlick claimed that a historical survey shows that even the earliest thinkers, without explicitly stating it, have still had exactly the same concept of metaphysical knowledge; however, Schlick did not provide such a survey or refer to one, so this was a completely unsubstantiated claim.

I show in Subsection 3.1.3 of this work, Newman has shown that our knowledge of the transcendent even in the special sciences cannot be purely structural, so this undermines part of Schlick’s reasons of resisting the act-object analysis.

32 What is extremely strange is that Schlick had previously, in the pre-positivistic phase of his philosophy, known that this was wrong. He had said in an article translated
As Ayer made clear (in [Aye34, page 386]), the attack of logical positivists on metaphysics was far more radical than that of Kant; logical positivists held metaphysical questions to be not only in principle beyond our knowledge (as Kant did the questions of transcendent metaphysics) but actually meaningless. Nevertheless, the attack is a continuation and radicalization of Kant’s attack, and perpetuates Kant’s misunderstanding of the nature of metaphysical enquiry.

Actually Ayer’s attack turns out in the longer version to have other premises than verificationism. Surprisingly, he added in [Aye36a, page 24] that the fact that the utterances of the metaphysician are nonsensical does not follow simply from the fact that they are devoid of factual content. He said that it follows from that fact together with the fact that they are not a priori propositions. This is very surprising, since Kant, to whom Ayer so often refers, claimed that metaphysics was a priori according to its very concept, and many metaphysicians of Ayer’s time also thought metaphysics to be a priori. Why did Ayer think that metaphysical statements are not a priori propositions? The answer is surprisingly obscure. Ayer referred to his arguments in [Aye36a, chapter 4] for the view that a priori propositions, including in his view all mathematical statements, are in his view analytical, and this means that they are tautologies. Ayer did not think that analytical propositions would be meaningless and this was important for him, since he thought that all correct philosophical statements, of course including those that he himself made, are analytical.

Why then did Ayer think that metaphysical claims cannot be analytical? Unfortunately, so far as I can find out, Ayer never told us this. Because Ayer did not demonstrate that metaphysical statements cannot ever be analytical, his demonstration of the impossibility of metaphysics fails utterly. As we have seen Aristotle, the very founder of metaphysics, thought that some in [Sch79c, page 110] that the task of metaphysics is a matter of so fashioning and coordinating the results of the special sciences as to produce from them a closed, harmoniously complete world-view. This is quite close to the original Aristotelian and the Russellian definitions of metaphysics. Schlick had then explicitly denied that metaphysics would require to employ other and more speculative methods than the rest of the sciences. Somehow he had under the influence of Wittgenstein and other logical positivists forgotten this. However, it must be admitted that Schlick’s earlier view of metaphysics was not based on historical evidence any more than his later one, though as I have argued historical evidence shows it to have been more correct.

33In [Aye87, page 32] Ayer gave up the view that these statements are tautologies, but still held that all of them are analytical.
metaphysical statements were such as Ayer would have thought analytical, so the view that some metaphysical statements are analytical has strong historical backing. It is indeed highly implausible on the face of it that all metaphysical statements would be tautologies, but scarcely more implausible than Ayer’s claim that all mathematical statements are tautologies, so if Ayer could swallow the later claim, why not the former?\footnote{There is a deep-going problem in the way Ayer (and most logical positivists) conceived of analytical statements. Ayer held (see e. g. [Aye36a, page 71]) that all mathematical truths were analytical, but Ayer also said in [Aye36a, page 128] that to assert that an object exists is always to assert a synthetic proposition. However, mathematics most certainly contains true existential assertions - e. g. the claim that there is a prime number between six and nine. This produces a contradiction in Ayer’s theory. This shows that analytical statements cannot be as trivial as logical positivists thought them to be and yet have so many uses as they thought them to have. This contradiction does not occur in most of the theories of most earlier non-positivistic analytical philosophers; e. g. Frege nowhere denied that analytical truths could be existential and though Russell claimed that mathematics was reducible to logic (a question on which Ayer suspended judgment), he claimed initially in [Rus03, §434, page 457], contrary not only to Ayer’s but to Kant’s and even Frege’s opinion, that logic is synthetic.} Ayer seems to accept blindly Kant’s claim that metaphysics must be synthetic while rejecting Kant’s claim that it must be a priori (or perhaps he just runs together the concepts of empirical and synthetic since he thinks them to be coextensive). This is made especially perplexing by the fact that several times in the course of his exposition of analysis, Ayer said that many statements which have been thought to be metaphysical (such as Berkeley’s claims) are not really metaphysical but are really analytical and hence tautologies. Could he not just as well have said that they were really metaphysical, but many metaphysical statements which had been thought to be synthetic were really tautologies?

For example, Ayer argued quite plausibly in [Aye36a, pages 38-39] that Berkeley’s immaterialist theory resulted from his analysis of the notion of a material thing (an analysis which Ayer himself actually held to be at least partially correct). However, where Ayer inferred from this the absurd claim that Berkeley’s theory was not really metaphysical, he could as well have inferred that Berkeley’s metaphysical view resulted from Berkeley’s employment of the method of analysis. This would imply that at least a vast part of Berkeley’s philosophy (roughly that which remains if Berkeley’s theology is removed from it) is an example of a metaphysical theory consisting of statements that are analytically true or (as those who unlike Ayer disagree with Berkeley’s analysis must claim, if they accept the very notion of analyticity)
analytically false.

In fact the general philosophical position Ayer arrives at by the method of analysis contains many claims that are naturally counted as metaphysical. Ayer held (see [Aye36a, page 130]) - as also many other early logical positivists (e.g. Schlick as seen in [Sch32b, page 5]) held - that both material and mental entities could be analysed in terms of sense-contents which were neither mental nor physical, more specifically regarded as logical constructions out of them (i.e. defined implicitly by means of them). This is a metaphysical doctrine that is known as neutral monism. Leopold Stubenberg has given a good introduction to this theory and its history in [Stu10]; according to him David Hume (and possibly Spinoza) can already be argued to have anticipated this theory, but it was first explicitly formulated by such positivists as Ernst Mach (in [Mac97], Richard Avenarius and Joseph Petzoldt. It had also been held by earlier British analytical philosophers (who arguably were not positivists) like Bertrand Russell in [Rus18a] and in [Rus18c]. Russell, however, explicitly described his theory as metaphysical, saying in [Rus18a, page 125] that he wished to discuss an ancient metaphysical query. Schlick actually referred to Avenarius and Mach in [Sch32b, page 5] (which was a source of Ayer’s view) and argued that the view that sense-contents are neutral distinguishes positivism from idealism (which held them to be mental or subjective) as well as realism, and tried to argue on the basis of this that positivism is not a metaphysical theory. However, all that this argument could show would be that positivism is a third metaphysical theory beside realism and idealism, not that it is not a metaphysical theory at all. Leopold leaves no doubt that in his

Schlick also argued that the founders of positivism certainly wanted something quite different from a renewal of idealism, and hence an interpretation of positivism as idealistic should be rejected as incompatible with the anti-metaphysical attitude of positivism. However, that the founders desired something different form idealism certainly does not suffice to show that they succeeded in founding something different from idealism.

Whether it shows even this is a controversial and difficult question. Neutral monism was introduced as a third position between materialism and idealism, but realism is not the same as materialism. However, it is controversial whether neutral monists managed even to stake out a third position between idealism and materialism; it can be argued that most early logical positivists were implicitly idealists, while later ones were implicitly or even explicitly materialists. Many philosophers have argued that neutral monism is idealistic, e.g. Lenin in [Len47]; however, Lenin was likely to label any theory other than his own orthodox Marxism as idealistic, so his testimony must be treated with care. Actually the very question is ambiguous and its answer depends partly on how the terms occurring in it are understood, i.e. on how realism, idealism and materialism are defined, and partly on the kind of neutral monism in question. Ayer took idealism to hold that
view this is a metaphysical doctrine; he says:

Neutral monism is a monistic metaphysics. It holds that ultimate reality is all of one kind.

It can be held to be an example of the kind of metaphysics that Kant thought to be possible, immanent metaphysics, since it is a doctrine about what can possibly be experienced.

Later logical positivists under the influence of Otto Neurath\textsuperscript{37} often everything that exists (in any sense of the word "exist", which Ayer took to have several meanings) is mental (see e. g. [\textit{Aye36a}, page 151], and if idealism is understood in this strong sense then his view might indeed count as neither idealistic nor realistic. However, idealism can also quite naturally be understood in a weaker sense so that not only a view according to which everything is mental, but also a view according to which everything ontologically fundamental is mental or depends on what is mental, so that for it to exist is either to perceive or be perceived, counts as idealistic. Since Ayer said in [\textit{Aye36a}, page 153] following Berkeley that though sense contents (which as we have seen were ontologically fundamental in his view) were not themselves mental, yet for sense contents to be is to be perceived, this suffices to class his view as definitely idealistic. Also since in Ayer’s view for sense-contents to be is for them to be perceived by individual human consciousnesses, rather than some kind of cosmic consciousness, this seems to suffice to classify his view even as a representative of subjective idealism. However, since Russell admitted in [\textit{Rus18c}] that sensibilia could exist without being sensed, his view can be considered realistic in a weak sense (though not materialistic). It certainly presents an alternative to traditional kinds of materialism and idealism, but it definitely presents a third metaphysical theory. Its view of a six-dimensional world, analogous (as Russell admits in [\textit{Rus18c}, page 158]) to Leibniz’s monadology, is definitely a piece of speculative metaphysics, though a quite well justified piece given the controversial theory of experience on which it is based; however, of course if the sense-datum theory of experience on which it is based is rejected, it must also be rejected. Therefore at least neither of these two kinds of neutral monism constitutes a third option between realism and idealism understood in the widest senses, though Russell’s theory constitutes a novel variety of realism in that wide sense. However, Carnap’s theory in the \textit{Aufbau} might constitute such a third option. Carnap holds that not only are physical objects reducible to mental ones, but the reverse is also the case (if epistemic considerations are excluded). This really implies that neither physical objects nor mental ones are metaphysically primary, and this does seem to be novel view (though not a very plausible one, as the justification of both reductions is highly questionable; as I will show in Section 4.3.1 that the very conception of reduction on which Carnap relies is dubious and thought to be incorrect by many philosophers today.).

\textsuperscript{37}It is often said that Neurath supported physicalism because he was a Marxist. There are indeed reasons to explain Neurath’s motivations in this way. Other Marxists, such as Lenin in [\textit{Len47}] had already argued against neutral monism on the ground that it was incompatible with Marxism. Lenin spoke of empiriocriticism, which was Avenarius’s term for the theory in question, as well as of Machism, and as we have already seen neutral monism was first explicitly formulated by Mach and Avenarius. There were many (today mostly forgotten) Russian neutral monists such as Bazarov, Bogdanov (remembered as the author of the first Bolshevik Utopia \textit{Red Star} and as a pioneer of blood transfusion), Yushkevich, Valentinov, Chernov and others Lenin primarily argued against. However, reference to Marxism does not explain Neurath’s position completely. Neurath’s physicalism was scarcely a version of dialectical materialism. It is uncertain if orthodox Marxists
abandoned neutral monism in favour of physicalism. Physicalism, however, is quite clearly an alternative metaphysical doctrine; it either says that all entities are physical (and so is an instance of metaphysics as ontology) or that fundamental entities are physical (and is an instance of metaphysics as aetiology). Clearly this constituted a gigantic alteration of metaphysical theory, though this of course took place without the positivists admitting that they had changed metaphysical doctrines or even without them always admitting that they had changed their opinions radically.

This later, physicalistic form of logical positivism is indeed unlike the early neutral monistic form of it inconsistent with one of Aristotle’s definitions of metaphysics, namely Aristotle’s definition of it as the science which deals with that which exists separately (from matter), as it holds that nothing exists separately from matter. It would indeed be consistent with metaphysics in that sense, if it were reformulated in a presupposition-free form so that it asked whether there was anything separate from matter and if so what, as I proposed earlier, for it would give an answer to that question: no, there is not. However, even if this reformulation is taken as too great a departure from Aristotle’s original definition, even physicalistic logical positivism is yet consistent with the legitimacy of metaphysics in all the other traditional senses, as ontology and aetiology and ousiology, most importantly with metaphysics as the science of being as being, which I have argued to be the most fundamental sense of the word.

The same obviously holds generally, even beyond the limits of the problem of realism; questions about the metaphysical status of material things such as Lenin would have admitted Neurath to be a materialist at all any more than they acknowledged Bazarov, Bogdanov and the rest as materialists. While Neurath’s theory may indeed be the better for abandoning the dialectical (in the peculiar Hegelian sense of dialectics) elements of Lenin’s materialism, yet there are elements in Lenin’s theory besides the dialectical woo-woo that may appear preferable to Neurath’s to many of today’s scientific realists. Lenin’s theory of reflection was a form of correspondence theory - not indeed explicitly of truth as such but of knowledge, but as knowledge is generally held to involve truth it suggests a correspondence theory for truth also. Such a correspondence theory is of course in radical conflict with the coherentism Schlick ascribed - in my view correctly - to Neurath. Neurath, on the other hand, would have accused Lenin’s theory of reflection of being metaphysical - and quite correctly too, though Lenin might not have agreed - so Marxism does not explain Neurath’s position.

This did not necessarily and perhaps not even usually imply acceptance of scientific realism; physicalistic logical positivists held that middle-sized physical objects such as rocks or trees (which would commonly be thought observable, though sense-data theorists did not think them such) existed and were not mere logical constructions, but did not generally hold the same view of unobservable physical entities such as quarks or electrons.
and minds appear in logical positivism (and more generally in analytical philosophy) as questions about the correct analysis of statements purportedly concerning material things and minds. Even more generally, questions about the metaphysical status of any kind of entity appears as questions about the correct analysis of statements apparently speaking about entities of that kind. Metaphysical priority appears in logical positivism (and analytical philosophy generally) as analyzability (which is a form of reducibility). Entities of one kind are according to them metaphysically prior to those of another kind if and only if the entities of the later kind can be analysed as either a subclass of entities of the former kind (if they can be defined explicitly with the aid of them) or as logical constructions out of entities of the former kind (if they can only be defined implicitly with their aid) but not conversely.

It seems to me that this is a quite legitimate and important way of showing that some entities are metaphysically prior to others; however, it need not be the only legitimate way. While statements of metaphysical priority can be analytic, they can yet also be synthetic a posteriori. A philosopher who does not share Ayer’s metaphilosophical opinion that all philosophical statements are analytical can claim that another way to show that entities of one kind are metaphysically prior to those of another and those entities of the second kind are reducible to those of the first kind is to argue (inductively or abductively) that it is a posteriori necessary (in the way that Kripke has argued in [Kri72] that many statements are) that entities of the later kind are identical with entities of the former kind but not conversely.

One of the logical positivists, Gustav Bergmann, already noticed that the views of logical positivists (as well as those of earlier positivists) contained a metaphysics, and wrote an article with the title "The Positivistic Metaphysics of Consciousness" in [Ber45] and a whole book of his writings had the title "The Metaphysics of Logical Positivism" in [Ber54]. Bergmann explained in [Ber45, page 194] that by the choice of such a title he wanted to emphasize the continuity between positivistic and non-positivistic thought. The problem with this explanation, of course, is that logical positivists claimed that their theory was radically different from earlier thought so far as it was metaphysical, so if continuity between positivistic theories and metaphysical ones existed to be emphasized, logical positivists were radi-
cally wrong. Therefore instead of a mere difference of emphasis, there was implicit a radical substantive difference of doctrine between Bergmann and the rest of the logical positivists. Indeed, following his train of thought to its logical conclusion Bergmann later rejected logical positivism altogether (though he had been initially very positivistic indeed, criticizing Carnap for not being positivistic enough in [Ber44]) and became one of the classical 20th century metaphysicians. Bergmann can be said to have overcome logical positivism from within far more then Neurath, to whom Uebel ascribes such a feat in [Ueb92]. However, there are many metaphysicians who think that Bergmann was not sufficiently metaphysical but remained too much of a linguistic philosopher.

One of the main contributions of Bergmann to the development of philosophy was that he reminded philosophers that the artificial language method in philosophy was not inseparably connected to positivism (as current positivists such as Lutz still seem to think), but could be employed in metaphysics. However, some of Bergmann’s views about the use of artificial languages in metaphysics are curious and show that he had not completely freed himself from the influence of logical positivism (at least at the stage of his philosophy during which he wrote [Ber54]) and might also have been adversely influenced by ordinary language philosophy. Bergmann described in [Ber54, page 102] the difference between himself and Quine so that Quine wishes to reconstruct the old ontological assertions, or some reasonable equivalent of them, in the ideal language, while Bergmann wishes to explicate them in informal discourse about the ideal language. Ontological

\[\text{39} \text{Lutz claims in [Lut12] that the underlying methodology of artificial language philosophy or ideal language philosophy, the development of languages for specific purposes, leads to a conventionalist view of languages in general. However, the expressions “artificial language philosophy” and even more “ideal language philosophy” are best suited to designate the whole long and big tradition which started in modern philosophy from Frege (with antecedents in earlier thought in Leibniz, Wilkins and others), whose Begriffsschrift was surely a paradigmatic artificial language (in the philosophical sense, which is very different from the sense in which such auxiliary languages as Esperanto or such artistic languages as Quenya or Klingon are artificial languages). However, this tradition by no means leads necessarily to conventionalism. It is not essential to that approach to philosophical questions that languages could be developed only for specific purposes, but many of its proponents, including the father of the approach, Frege, believed in the possibility of a universal ideal language in which all questions could be addressed. The radical conventionalism into which one of the followers of the method, Carnap, arrived is by no means essential to the method. Most of the exponents of the method not affected by the positivist tradition would have rejected such conventionalism. Frege argued at length against Hilbert’s conventionalist theses.}\]
assertions are surely assertions about the world in general, about being as being, not about any language, and therefore the position Bergmann ascribed to Quine seems to me far preferable to Bergmann’s own position. I will compare Bergmann’s metaontological views further with Quine’s in 5.2.1.

3.3 Problems with Verificationism in General

We have seen that even a strong version of the verificationist criterion would not imply that metaphysics would be meaningless tout court, but at most that metaphysics would be impossible in one derivative sense of the word, namely in the sense of transcendent or speculative metaphysics, and that in other senses of the word a certain kind of metaphysical theory, such as subjective idealism or neutral monism, would be correct. Should we, however, agree that metaphysics is impossible even in that one derivative sense, as knowledge of the transcendent? Have we in responding to the logical positivistic attack on metaphysics found a method of metaphysical research, ending up with a constructive rather than a merely negative result? Do we have any reason to think that the verificationist criterion gives us any reason to think that such an anti-realistic metaphysical theory is true (and that more realistic theories are not)?

In this subsection of my dissertation I will argue that the answers to these questions are mostly negative, since the verificationist criterion itself faces very serious problems. Because of them logical positivists did not succeed in giving any good reasons to think that even metaphysics in the derivative sense of knowledge of the transcendent, i. e. speculative metaphysics, would be impossible. In responding to verificationist attacks on metaphysics we may indeed have found one positive method. The method of conceptual or meaning analysis may be a valid method of metaphysical research; however, it is not a valid method if it is used in the way verificationists used it, but only if it is used as non-positivistic analytical philosophers used it.

Because of the verificationist principle metaphysical priority was in Ayer and the rest of the logical positivists equated with epistemological priority (so that they ironically arrived in this question at a similar result as Neo-Aristotelians like Simons or Tahko who in other ways oppose them). However, the method of analysis itself need not lead to this result if it is separated from verificationism. If meaning consists (even partly) in verification con-
ditions or assertibility conditions or confirmation conditions etc., as most serious forms of verificationism claim, then meaning analysis will naturally end in entities which (or the experience of which) would ultimately verify sentences, and in the case of actually justified sentences entities (the experiences of) which actually do ultimately verify them, i.e. epistemologically primary entities. However, if meaning instead consists for example solely in truth conditions, then meaning analysis need not end in anything epistemologically basic. The fully analysed statements containing only primitive terms which are the ultimate results of meaning analysis do not in this case have to speak about observable states of affairs. Even if the basic terms that occur in them would have to denote observable properties or particulars, as the empiricist Principle of Acquaintance implies, the properties and particulars can in such fully analysed statements be combined into unobservable compounds. Moreover, the fully analysed statements can contain existential generalizations such that the bound variables occurring in them have as values unobservable entities. Of course, it is highly controversial whether Ayer’s analysis of the notion of material objects is correct; most philosophers today would think that it is not, and I am inclined to agree, but the question is not as easy to answer as it is often supposed to be. Those who reject verificationism are indeed far less likely to agree with such an analysis than those who accept verificationism. However, I suppose it is in principle possible to consistently accept such an analysis without accepting verificationism; Russell accepted such an analysis in [Rus18c] without accepting full-blown verificationism, though he accepted a similar maxim.

While the strong version of the verificationist criterion would not outlaw all metaphysics, it would outlaw many statements of natural sciences which the logical positivists would not have wanted to abandon. It can be argued that even forms of empiricism far weaker than verificationism are not compatible with a scientific world-view based upon a correct interpre-

\[40\] This supposes, of course, that existential generalizations cannot be analysed as disjunctions of atomic statements, as e.g. Sellars proposed in [Sel49]. If such an analysis were correct, then the Principle of Acquaintance would imply that meaning analysis would have to end in epistemologically basic entities. Some logical atomists may have gone so far as to accept such an analysis. However, most logicians today agree that such an equivalence is not correct. Actually this kind of analysis would imply that all quantification would be substitutional, so if objectual quantification is a logical notion, existential generalizations can concern unobservable entities. I will later in this work deal at length with the distinction between substitutional and objectual quantification.
tation of modern science, since mathematics is essential to modern science and empiricism has always found it very hard to account for our knowledge of mathematics. The most sustained attempt to deal with this problem was logicism, which as we have already seen claimed that all mathematical statements were analytical and indeed logical; however, since logicism has encountered severe difficulties, this threatens the prospect of a reconciliation between empiricism and mathematics. In any case, even if weaker forms of empiricism could somehow surmount this problem (in one of the innumerable ways empiricist philosophers have developed to deal with it), verificationism poses new and far more difficult problems.

The greatest problem with logical positivism concerns the formulation of the verificationist criterion. The verificationist criterion can be formulated in different ways, in stronger and weaker versions. The lengthy debate about the principle showed that stronger versions lead to absurd conclusions which are incompatible with a scientific world-view and which therefore

41 There is an even worse problem, but it seems to afflict only some versions of logical positivism. It has often been argued (e.g. by A. C. Ewing in [Ewi37]) that strong formulations of the principle are self-defeating, since they imply that they themselves are meaningless. It is not clear that this is a valid criticism of all logical positivistic theories. At least those logical positivists who restricted the principle so that analytical statements were not meaningless could argue that the principle itself was analytical, though it was indeed unnatural to view it as a tautology. Most logical positivists did not regard analytical statements as meaningless, though there were a few early incautious formulations that might be taken to imply that they were such. E.g. Schlick claims in [Sch32b, page 21] (translated in [Sch49b, pages 98-99]) that a proposition has meaning only if I can state the conditions under which it would be true and the conditions under which it would be false. Since for an analytical statement there are no conditions under which it would be false, this formulation implies that all such statements, the verificationist principle included, are meaningless. Such formulations of the principle are then rather obviously self-defeating. It seems to me that Wittgenstein’s formulation of a principle about what can be said (which is similar to a principle about what is meaningful) in the Tractatus [Wit21] is almost defiantly self-defeating, as Wittgenstein himself flouts it constantly, trying to cover up this inconsistency with obscure metaphors about ladders. However, most formulations of the principle are not like this. Indeed, the strong form of the verificationist principle actually implied that it itself was analytically true, for if the meaning of every synthetic sentence is the same as a method of verifying it, then it follows that a synthetic sentence which cannot be verified cannot be meaningful. The strong form of the principle is then self-consistent, though unacceptable for other reasons, for instance because of being inconsistent with the scientific world-view it was supposed to be used in defending. However, it is not clear how it could be shown that weaker versions of the principle, which may be consistent with a scientific world-view, are analytic, so there is some reason to suspect that weaker forms of the principle might be self-defeating; nevertheless, they are not obviously self-defeating, so I will not rely on this argument.

42 On the other hand, these results are not compatible with a religious world-view either. While an objective idealism might be well suited for at least many kinds of religious world-view, I have argued that the implicit metaphysics of early logical positivists like Ayer was
the logical positivists themselves were not willing to admit, while the weaker
versions threaten to be completely vacuous, so that they cannot be used to
discredit metaphysics or indeed cannot be used for anything at all. The logi-
cal positivists themselves found that they had to constantly back away from
their initial position. It was shown that not even statements that positivists
thought fully legitimate and important fulfil the verificationist criterion, at
least in its stronger, original form.\footnote{The retreat of logical positivists from
the strongest forms of the criterion was so swift that there are few writings published
by the logical positivists themselves in which these strongest forms can be found;
the very strongest expressions of verificationism have been found by historians of
philosophy in unpublished manuscripts (which have of course often been
published posthumously by such historians). Ayer told us in two footnotes in \cite[
page 23]{Ayer34} that his views regarding the correct form of verificationism had altered
between the writing and publication of the article (which of course raises suspicions
about the consistency of the article together with its footnotes). From the view that
all meaningful sentences would have to be conclusively verifiable he had shifted to
the other extreme, now claiming that no sentences are conclusively verifiable. How-
ever, there are some published examples of strong verificationism; Carnap in the
second section of \cite[pages 221-224]{Carn32b} seems to have held a quite strong
verificationism.}

Of course different logical positivists differed in what consequences of
strong forms of the principle they found unacceptable, but most found some
of them to be such. Most logical positivists had a firm confidence in natural
sciences and many of them also trusted in in common sense, so they did not
want their attack on metaphysics to threaten them. However, they found
to their dismay that non-vacuous verificationism was just as opposed to
natural science (and less importantly common sense) as to traditional meta-
physics, implying that many statements of science and common sense were
meaningless. Such statements were statements of science concerning theo-
etrical entities such as atoms or subatomic particles (which the old positivist
Mach had already famously attacked). Subatomic particles can already be
said to underlie and transcend phenomena in one rather strong sense, since
they cannot be directly observed (at least without the aid of instruments),
and it is not clear that most metaphysicians require entities that would be
transcendent to phenomena in any stronger sense. Another example were
statements concerning the past. Another little known but very striking ex-
ample are general laws such as laws of nature, including in the case of truly
strong forms of verificationism even general laws stated purely in observa-

\footnote{A subjective idealism, and that kind of idealism does not jibe with a religious
world-view. Such a subjective idealism has no room for material bodies, which makes it a bad fit
for a scientific world-view, but neither has it room for a transcendent God, which makes it
also a bad fit for a religious world-view.}
tional terms. This consequence of very strong forms of verificationism would indeed be fatal to even immanent metaphysics, as generality is essential to metaphysics. However, in denying the possibility of the absolute generality required by metaphysics, strong verificationism also denies the possibility of the restricted generality required by all science (see for example [Aye34, page 340]).

Ayer originally tried in [Aye36a, page 20] to weaken the verificationist criterion so that a proposition (by which Ayer just meant a sentence) is empirically meaningful if some experiential propositions can be deduced from it in conjunction with certain other premises without being deducible from those premises alone. However, it was pointed out by Isaiah Berlin in [Ber38] that this formulation allowed any sentence to be meaningful. If $S$ is any sentence that the logical positivists would not want to be meaningful e. g. "The Absolute is perfect." and $O$ any experiential sentence we could just take $O$ if $S, S \rightarrow O$ as the other premise. Ayer then tried to find more complex formulations that would escape this trivialization, and Carnap attempted the same in [Car56], but such authors as Alonzo Church (in [Chu49]), David Kaplan, David Lewis (in [Lew88]) etc. have produced arguments showing that these new formulations shared the same problem. Authors such as Richard Creath (in [Cre76]) M. Przelecki, Sebastian Lutz (in [Lut10]) etc. have tried to reformulate the criterion so as to escape these arguments.

However, Creath does not claim that his formulation of the criterion

\footnote{The crucial idea in Lutz's attempts to escape trivialization of the verification criterion is that of honest sets of auxiliary sentences. A sentence is according to Lutz empirically significant when it implies observational sentences together with a honest set of auxiliary sentences. A set of sentences is said to be honest (see [Lut10, Definition 37, page 40]) when every sentence in it is justified and for every sentence in it, the set also contains every sentence on which its justification depends. If I have understood this definition correctly, there are many problems in it. Different sentences are justified for different persons and for the same person at different times, so a sentence could be empirically meaningful for some persons and not for others. Certainly most scientific theories would not be empirically significant for most people, and quite likely many scientific theories would not be empirically significant for any people, though they might become significant in the future. These results seem absurd, and perhaps more tellingly, just as harmful for science as the original formulations of the verification principle, for if an empirical science ought to reject empirically non-significant theories (since this is what the criterion is meant for) then Lutz's definition would lead us to reject theories that would be shown to be not only meaningful but even true in the future, as auxiliary statements needed to make them meaningful would only become justified in the future). If on the other hand talk of justification in this definition were replaced with talk about justifiability, the threat that the definition has become vacuous would surface once again.}
would provide a completely adequate criterion of empirical meaningfulness, and gives as a reason in [Cre76, page 399] that the criterion presupposes a classically sharp distinction between the observation and theoretical languages which Creath does not accept. It is essential to the originally intended use of the criterion to declare some sentences meaningless that the distinction between observation sentences and other sentences be sharp. E. g. Schlick clearly said in [Sch32b, page 9] that the distinction between the essential impossibility of verification and mere empirical impossibility was supposed to be perfectly sharp (scharf), since it is a logical one and differs from empirical not by degrees, but absolutely. If the distinction between observation language and theoretical language were not sharp, then obviously the distinction between the essential impossibility and empirical impossibility of verification could not be sharp either.

If all observation were indeed theory-laden, as those who deny the sharpness of the distinction usually think, then a proponent of a metaphysical theory could just observe it to be true. For example, a Leibnizian would see (with his eyes) physical objects as complexes of monads, since what he sees would be influenced by his beliefs, while a Spinozist would similarly see Nature as God. This would suffice to show that monads and God are observable and therefore that Leibniz’s and Spinoza’s theories are meaningful and even justified. This is surely too easy a way to even show a metaphysical theory to be meaningful, much less justified, even for one who believes some metaphysical theories to be meaningful and justified. Since Creath apparently denies the existence of a sharp distinction between observation sentences and other sentences, this is already a reason to claim that his theory gives no support to verificationism at all. However, I would not want to rely on this to support the possibility of metaphysics, for I would myself be very wary of denying a sharp distinction between observation sentences and other sentences as Creath does, since such a distinction is essential not only to verificationism but also to far weaker and more plausible kinds of empiricism. Lutz at least thinks (see [Lut10, page 5]) that a sharp distinction between observation terms and theoretical terms can be established, even if not in ordinary language then at least in artificial languages, so his support of verificationism (or at least of empiricism) is more full-hearted than that of Creath. The existence of a sharp distinction between observation terms and theoretical terms (or possibly of several sharp distinctions, if the word “ob-
servation” is ambiguous) is then necessary for any non-trivial empiricism\textsuperscript{45}; however, as I will argue below, it is by no means obvious how and where the distinction is to be drawn, and this is already a non-trivial problem in epistemology.

There is one ambiguity in the distinction pointed out for example by Hilary Putnam in \cite{Put62} and by Frederick Suppe in \cite{Sup89, pages 58-60} which I must dwell on for a while because it will become useful later in this dissertation. Is a property observable if its presence is sometimes ascertainable by direct observation or only if it is always so ascertainable? Analogously, to put the distinction in a way which even a nominalist who rejects properties might accept, is a predicate term observational if some of the atomic sentences whose predicate it is (and whose subject is an observational singular term) are observation sentences or only if all atomic sentences whose predicate it is (and whose subject is observational) are observation sentences? The distinction between observational and theoretical was originally a distinction between different sentences (or states of affairs or propositions), namely between sentences which were implied by observation sentences aka protocol sentences and other sentences which were not so implied (or between observable states of affairs and other states of affairs), not between different terms or entities in general. Therefore extending it into the latter domain involves complications which were not always noticed. Rather than a dichotomy between observable and non-observable properties, there is a trichotomy; however, this does not imply that the distinctions between the three classes, entities which are always, sometimes and never directly observable, would not be sharp.

We could say that a predicate term which always applies to observable entities is strongly observational, while a predicate term which sometimes applies to observable entities is weakly observational. Similarly a predicate can be said to be strongly theoretical if it is not weakly observational and weakly theoretical if it is not strongly observational. However, the more important division seems to me to be between properties which are sometimes and those which are never observable or between weakly observational and

\textsuperscript{45} If perception, including observation, is held to be non-conceptual, as proposed by Christopher Peacocke in \cite{Pea01}, then obviously it need not be theory-laden. However, even if perception is held to have a conceptual content, it need not be held to be theory-laden, for the most primitive kinds of concepts can be held to exist apart from written language. I will deal with this debate at more length later.
strongly theoretical predicates. It can be argued that there are no strongly observational predicate terms (as Putnam did argue), in which case all predicate terms would be weakly theoretical; however, even if Putnam is wrong, as he seems to be, it seems clear that if there are strongly observational terms, they are rare indeed. Also a weakly observational term can be defined ostensively, so what Schlick at least thought to be the ultimate purport of the distinction is satisfied by a distinction between weakly observational and strongly theoretical terms. Because of this these can be called observational and theoretical terms simpliciter. The fact that some properties are sometimes but not always observable is indeed highly fortunate. This is the only reason that enables us to bridge the gulf between observable and theoretical entities and have some knowledge of unobservable entities, whether this is by defining theoretical terms by means of observational ones or using induction to get from observation statements to theoretical ones, as I will argue later in Section 6.1.

However, even if we set aside the problems concerning the distinction between observational and theoretical, the verificationist criterion faces serious problems. Responding in [Aye36b, page 199] to the critique of W. T. Stace in [Sta35], Ayer said that he did not require of a meaningful statement that it should be conclusively verifiable, but only that some possible observations should be relevant to the determination of its truth or falsehood. Ayer called this verifiability in a weak sense, but it might be less misleading to call it a disjunction of defeasible confirmability and disconfirmability. We can call this view weak verificationism, while the view which requires conclusive verifiability can be called strong verificationism. Ayer still asserted

46If observational terms are taken to apply to physical objects as the later logical positivists did, rather than to sense-data, as in the earliest versions of the distinction, then it is quite likely that there are no strongly observational terms in natural languages. Nevertheless, we can of course always artificially manufacture strongly observational predicates out of weakly observational ones by restricting their extension to observable particulars (which shows Putnam to be wrong), though such a procedure has little point if we are not verificationists or constructive empiricists.

47The distinction between weak verificationism and strong verificationism is obviously parallel to that between weak foundationalism and strong foundationalism, as made by e. g. Susan Haack in [Haack99, page 54], which we have already encountered. Weak verificationism implies weak foundationalism but not conversely, and similarly strong verificationism implies strong foundationalism but not conversely. Neither of weak verificationism and strong foundationalism implies the other. While I will argue that weak verificationism as a criterion of meaningfulness collapses into triviality, the same does not hold of weak foundationalism as a theory of justification.
confidently that this criterion denies meaningfulness to metaphysical statements such as those which deny the reality of the sensible world or assert the existence of things in themselves: however, there are good reasons to think that this is not correct.\footnote{In \cite{Aye67} and in \cite{Aye87} Ayer at last relented to some degree of his opposition to metaphysics; however, his admission of the legitimacy of metaphysics was very back-handed. Ayer said in \cite[page 33]{Aye87} that he did not abide by his summary dismissal of metaphysics because it had occurred to him that metaphysicians have suffered from the assumption that they literally meant what they said. Apparently Ayer still did not admit that any statements of metaphysicians are to be considered seriously if they are interpreted literally and not metaphorically. However, the reasons Ayer had for rejecting metaphysics if interpreted literally were obviously now entirely different. Ayer claimed in \cite[page 64]{Aye67} that it can hardly be disputed that metaphysics is nearly always in conflict with common sense. Now metaphysics could hardly be in conflict with common sense if it were meaningless; logical contradiction can only exist between meaningful sentences. However, again Ayer did not give any argument based on references for his view that metaphysics is nearly always in conflict with common sense; it seems that he still took such extreme idealists as Bradley as typical metaphysicians and ignored all the rest. Indeed, the theory of logical positivists can be argued to be at least as much in conflict with common sense than most professedly metaphysical theories. It is hardly correct that Aristotelian metaphysicians would be nearly always in conflict with common sense. They seem rather to me to suffer from exactly the opposite fault; they try to stick to common conceptions too far, even when they come into conflict with natural science, which has often made them an obstacle to the progress of science. Even modern metaphysicians can hold like Peter Strawson in \cite{Str59} that metaphysics is to be descriptive metaphysics, which merely tries to explicate the implications and presuppositions of common sense. However, even if metaphysics tries to go beyond common sense - as I think it must if it is not become completely trivial - it need not yet get into contradictions with common sense, as it can only try to give answers to questions with regard to which common sense is neutral.}

In order to show what happens when the principle is weakened, we must notice that Ayer did not apply the weakening of the criterion consistently to all questions. He admitted that statements did not have to be equivalent to statements about the evidence that might be presented for them in some cases but not in all. He admitted that statements about past events were not equivalent to statements about the present evidence from them, but still assumed that statements about the mental states of other people were equivalent to statements about the behavioural evidence for them (so committing himself to logical behaviourism) and that statements about physical objects were equivalent to the statements about sense data that in his view provided the evidence for them. However, all three cases are similar and should be treated in the same way, at least unless some additional argument is given for treating them differently.\footnote{Similarly to Ayer, Schlick said in \cite[page 345]{Sch36} that it is quite irrelevant whether verification can take place in the future or in the present only. Why then is it relevant} There might of course be such valid
arguments, as many other philosophers have provided arguments trying to show a relevant difference between the cases. However, Ayer did not provide or even refer to any such argument and therefore even if some such arguments would be valid, Ayer would not have been justified in treating the three cases differently. Intuitively, we might perhaps even say according to common sense, the mental states of a person are causes of his behaviour and therefore serve to explain his behaviour just as past events are causes of present events and serve to explain them.

In [Aye87, page 31], Ayer finally admitted that his behaviourist treatment of the mental states of others was not justified by the weaker version of verificationism he adopted. Carnap likewise quite explicitly abandoned behaviourism in [Car56, §XI, pages 69-75]. Also other philosophers within or associated with the logical positivist movement proposed more plausible forms of materialism than behaviourism, such as the identity theory introduced within the movement by Herbert Feigl in [Fei58] and outside it by U. T. Place (in [Pla56]) and J. J. C. Smart. Unfortunately, such admissions have not been very influential, and there have still been recently many philosophers who try to justify behaviourism with appeal to some seldom clarified version of verificationism. This apparently includes even such influential philosophers as and Quine, whose argument for ontological relativity in [Qui68] has behaviourism as an essential and indeed explicit premise, as seen in [Qui68, page 186]. In my view this also includes Michael Dummett and his followers, as I will argue later, though unlike Quine Dummett denied in [Dum78, pages xxxii-xxxvii] that his theory, or the theory of Wittgenstein it is inspired by, is behaviourist. In my view this already suffices to undermine both Dummett’s justificationism and Quine’s ontological rel-

whether it can take place in the experience of all subjects or in the experience of one subject only? Why should the heteropsychological be any more problematic than the heterotemporal? Indeed, if all that is required is the logical possibility of verification, as Schlick claimed, then verification by means of telepathy surely suffices for meaningfulness, for telepathy is surely logically possible.

50Dummet dealt with the example of pain-ascriptions, which were used by Strawson as arguments against him. Dummett holds that Wittgenstein’s and his theory is neither behaviourist nor realist, but it is not clear that there is a middle way here; even materialist theories competing with behaviourism, such as the identity theory or functionalism, already seem to be realist theories. In order to refute realism about pains Dummett appeals dogmatically to Wittgenstein’s private language argument, not even bothering to repeat the argument, which is rather high-handed, since there is a lot of unclarity about just what that argument is supposed to be and where its probative force is supposed to consist.
ativism.

It is indeed not wholly clear why Quine embraces behaviourism. Quine seems to think that naturalism requires behaviourism; however, it is far from clear just what Quine means by "naturalism" in this context. The word "naturalism" is ambiguous in many ways, and can be taken either as an ontological or an epistemological (or methodological) theory (or some kind of combination of both). If naturalism is taken as an ontological theory, then the strongest form of naturalism is commonly taken to be physicalism (which Quine indeed explicitly accepted in many of his later writings, as in [Qui93]) or physicalism together with nominalism (which Quine at first defended but later abandoned). However, even physicalism does not require behaviourism, but is consistent with many different theories incompatible from it. It is generally held to be consistent (though there is much dispute about all of these examples) with identity theory (like Feigl's theory), functionalism (as many philosophers have recently argued, e. g. Devitt in [Dev84, page 277] arguing specifically against Dummett), emergent materialism, anomalous monism etc. While all of these theories have severe problems, all of them are generally recognized to be better than behaviourism. Neither, however, does moderate empiricism, which is often called epistemological naturalism, lead to behaviourism. Therefore one must suspect that some kind of verificationism or other kind of extreme empiricism (which would be extreme epistemological naturalism) remains as a motive for Quine's behaviourism. Since the ontological relativism at which Quine arrives as a result of his argument based on behaviourism is scarcely consistent with any kind of genuine physicalism or even weaker forms of naturalism (as a genuine physicalist would say that physicalism is absolutely true, not just true relative to some conceptual scheme, and similarly even a weaker kind of ontological naturalist would say that naturalism is absolutely true) then if Quine's argument is valid, surely even a physicalist is best advised to steer clear from behaviourism.

The situation with regard to physical objects is not quite the same as with regard to the mental states of other people or the past, since it can doubted whether the existence of sense-contents or sense-data is at all part of common sense; however, other sense-data theorists had claimed that physical objects were causes of sense data but could not be analysed into them, and Ayer did not really have any good arguments against this view. Ayer
had admitted that statements about the behaviour of other people can be relevant to the determination of the truth or falsehood of statements about their mental states (without being conclusive evidence for such statements) even if statements about mental states are not analysed in terms of such evidence. Similarly statements about sense data (or about sensations) could be relevant to the determination of the truth or falsehood of statements about physical objects whether or not physical objects are reducible to sense data. Or course it is only in the first case that they might provide conclusive verification, and in the second case they can only provide confirmation, but Ayer had admitted that conclusive verification is not needed. Therefore the weakened verificationist principle does not by itself decide whether or not physical objects are reducible to sense data, and therefore this remains a genuine question, and it is clearly very closely connected with the traditional metaphysical question about the existence of the external world. A representationalist could argue that the coherence of our sensations confirms the hypothesis of an external world, since if such a world did not exist our sensations would be unlikely to be organized in such systematic ways as they actually are.

However, it would not be fatal to the defender of metaphysics even if the particular statements Ayer mentioned were not verifiable even in the weak sense in which he now understood verifiability. In order for the principle to show that all metaphysics is impossible, all metaphysical theories would have to be such that no observations would be relevant to the determination of its truth or falsehood, but this is clearly not the case. A defender of metaphysics could easily accept that some metaphysical theories, such as some of the most extravagant specimens of German idealism (like the statements of Bradley which Ayer quotes), were meaningless; he has only to defend the thesis that some metaphysical theory (which is of some interest) is meaningful. It is of course not implausible that some theories which were not metaphysical had been misclassified as metaphysical; however, the positivist enemy of metaphysics would have to show at least that most statements that have been traditionally viewed as metaphysical or all statements that satisfy a traditionally used definition of metaphysical

---

51Surely it is logically possible that our sensations would suddenly lose all their coherence - that for example when we drew nearer to distant objects they would not seem larger but smaller, that when we walked round an object it would turn out to have no back side, etc. - and this would surely count against i. e. disconfirm the reality of the sensible world.
statements were meaningless. It is all too easy to just define metaphysical statements as meaningless ones, but if the positivist does this (as many have been accused of doing) he commits the most egregious possible form of the fallacy of attacking a straw man. If we look at one of the earliest highly controversial metaphysical theories, Aristotle’s theory of intelligences, we see that surely some observations are relevant to its truth or falsehood. Since the theory said that these intelligences move heavenly spheres and therefore the planets attached to them in a perfectly spherical motion, astronomical observations that indicated that the planets did not move in this way or even that spheres did not exist would count against its truth. Indeed, as we know from the history of science such astronomical observations were actually made by Copernicus and others and they led to the abandonment of the Ptolemaic astronomy and also to the abandonment of this metaphysical theory at least in its original form (though of course Neo-Aristotelians sometimes try to defend radically modified forms of this theory). Aristotle’s Intelligences were therefore not transcendent to experience in any stronger way than subatomic particles like electrons are.

The debate over the criterion has produced a long string of counterexamples and attempted patches, and most philosophers have ultimately lost interest, though the issue remains not fully decided and important. There is in this situation reason to suspect that no formulation of the verification criterion will be both acceptable and inconsistent with metaphysics, so that the obstacle to metaphysics that the criterion seemed to present was just an illusory one. Since the debate has rather petered out than been definitely resolved, it cannot be completely ruled out that some version of the criterion might ultimately be salvaged, unlikely though that seems at this point. Verificationism is not then quite as conclusively disproved as structuralism (in its original form) was disproved by Newman’s argument. However, there is no reason to think that such a criterion, even if one were found, would have the consequences logical positivists thought it would have, and would declare all metaphysics (or even a significant portion of contemporary metaphysics) meaningless. Besides, showing that the criterion could be formulated in a way that was neither absurd nor vacuous would not do

\[\text{I am not claiming that Aristotle’s arguments for the existence of the Intelligences would have been valid even if Eudoxus’s astronomy had been wholly correct; just how the arguments are supposed to work is rather murky. However, the vital point is that Aristotle did consider observations relevant to the adoption of his metaphysical theory.}\]
much to show that the criterion was true, for just showing that a statement is not absurd certainly does not suffice to justify that statement! The defenders of the criterion have lost sight of this crucial additional task of positively justifying the principle in the heat of responding to attacks on it.

Cheryl Misak says in [Mis95, page viii] that the thought at the heart of the verifiability principle is that a belief with no connection to experience is spurious. If verificationism were understood this broadly, then its roots would not only go as deep as Berkeley, as Misak says, but even Aristotle and such Aristotelians as Thomas Aquinas would count as verificationists (and probably even earlier philosophers, Democritus etc.). After all, Aristotle at least admitted that all beliefs must ultimately consist of primitive concepts derived from experience, which is a connection to experience, even if a rather weak one. However, this is absurd; the word ”verificationism” just cannot be appropriately used as broadly as this.

Indeed, Misak soon begins to argue for a rather strong theory, similar to the earlier version of verificationism that logical positivists themselves eventually abandoned. She argues that verificationism is inconsistent with realism, sliding soon from the claim that a belief must be connected to experience to the rather different claim that a belief cannot transcend experience, so it is rather pointless to begin from so weak a claim, unless it is done in order to mislead the reader.

Apart from the question whether verification has to be conclusive or not, there is another way in which formulations of verificationism can be ambiguous, namely the interpretation of the concept of possibility used when we speak of verifiability, i. e. of what can be verified, i. e. what it is possible to verify (or falsify or confirm or disconfirm). Ayer in [Aye34] had only spoken ambiguously of verifiability in principle, and this could be interpreted as either logical or nomological possibility of verification. Once logical positivists noticed the question, disagreements with regard to it arose among them. Moritz Schlick had formulated verificationism in terms of mere logi-

---

53 Yet in the end Misak’s lengthy discussion of various forms of verificationism leaves rather obscure just what kind of verifiability Misak requires. In the end she calls her view modest verificationism, and has such formulations as in [Mis95, page 201-202]

a statement lacks legitimacy or objectivity if there would be no evidence for or against it.

. This formulation does not even entail empiricism; even an extreme rationalist could agree with it, just adding the gloss that evidence need not be empirical evidence. Only a negative coherentist could disagree with it.
cal possibility of verification in 1936 in [Sch36]. On the other hand, Carnap held later in [Car36, page 423]) that the possibility in question was physical i. e. nomological possibility. Schlick’s view threatens to trivialize the principle completely. Schlick’s formulation of the principle was then in one respect rather weak: however, in another respect it was still rather strong, since Schlick still thought that it was the logical possibility of conclusive verification that was required.

Schlick even went so far as to say in [Sch36, page 342] that in stating the verificationist principle the positivist does no more than describe a situation in which there is no way of understanding any meaning without ultimate reference to ostensive definitions. However, this is clearly false; even the weaker form of verificationism goes far beyond the necessity of ostensive definitions. Even Aristotle could have accepted that necessity, since as we have seen he was a concept empiricist. The necessity of ostensive definitions was clearly expressed in Bertrand Russell’s Principle of Acquaintance as formulated e. g. in [Rus12b, page 91]:

> Every proposition which we can understand must composed wholly of constituents with which we are acquainted.

However, when Russell formulated the Principle of Acquaintance he yet accepted that we can have a lot of knowledge by description of entities with which we are not acquainted. Also the principle allows that there may be

---

54 This view is of course utterly incompatible with the structuralism also frequently held by logical positivists; e. g. with Carnap’s view in [Car67, §II.A.16, page 29] that anything that can be pointed out in a concrete ostensive definition is subjective, and since scientific statements are objective their meaning is independent of all ostensive definitions. Schlick himself had subscribed to Carnap’s view in [Sch26]. He seems then to have changed his view radically in ten years, unless his view was radically inconsistent.

55 We can here see an anticipation of Quine’s views in early Russell, as Quine recognized in [Qui53d, page 12]. Russell famously interpreted descriptions (both definite and indefinite) with the aid of quantifiers; so in having knowledge of an entity by description we are quantifying over it. i. e. the entity is a value of a bound variable, even though no basic constant term denotes it. Quine emphasized that the only way we can involve ourselves in ontological commitments is with the aid of variables. This is, however, very controversial, for opponents of descriptivism who follow Saul Kripke’s [Kri72] would say that names cannot be reduced to descriptions as both Russell and Quine thought. However, it seems to me that a more important point about variables is that we can involve ourselves with the aid of variables and descriptions in ontological commitments in which we cannot involve ourselves with any other means. It seems to me that even opponents of descriptivism ought to accept this very important fact. At least they ought to accept that descriptions play an important role in fixing the reference of names; by using descriptions to fix reference we can give references to names to which we could not otherwise have
propositions whose constituents we are not acquainted with; we just cannot understand such propositions. Russell thought at that time that we were not acquainted with physical objects (or with minds) but only with sense-data and mental acts (and their characteristics); however, he yet thought that physical objects (and minds) existed and we had knowledge of them. The verificationist principle in its early, strong forms denied the possibility of this, since according to it a sentence containing as its grammatical subject a definite description which purportedly denoted an entity which was not observable was meaningless. However, nothing in Russell’s theory implied that such a sentence would be meaningless, so verificationism is stronger than Russell’s Principle of Acquaintance. It is then important to distinguish these principles, especially since they have been commonly confused. Even if all our basic concepts and/or the semantic values of ba-

given it, even if the meaning of the description then falls away from the meaning of the name, unlike what Russell thought.  

56It is no part of the concept of acquaintance itself that we can only be acquainted with sense-data, as it has often been misunderstood. Russell’s view that we are not acquainted with physical objects did not follow from the definition of acquaintance, but from the Argument from Illusion, and philosophers who accept the Principle of Acquaintance can yet disagree whether that argument is valid. The Principle of Acquaintance is compatible with different phenomenological views about what we are acquainted with, and leaves room for theories according to which we are acquainted with physical objects, so that even a non-eliminativist physicalist might accept it. It seems very plausible to me that any theory which can be properly called concept empiristic (as distinct from empiricism concerning justification) must accept the Principle of Acquaintance if that principle is understood in the weakest sense, separated from accidental accretions. Indeed, the principle has sometimes also been called the Principle of Empiricism. The logical positivists’s concept of observation is historically derived from Russell’s concept of acquaintance and is certainly closely related to it if not the very same. However, while acceptance of the principle is arguably necessary for concept empiricism, it may not be sufficient; concept rationalists could also accept it, for intellectual intuition would also count as a form of acquaintance. An empiricist must suppose also that acquaintance happens in some way through the senses.  

57These principles seem to have been commonly confused in the ontological tradition started by Gustav Bergmann. Bergman and his followers thought that we had to be acquainted with the simples of ontological analysis. I suspect this view resulted (like Ayer’s similar view) from confusing epistemological and ontological priority; Russell’s original principle only says that we have to be acquainted with epistemologically primary i.e. epistemologically simple notions. E. g. Edwin B. Allaire appealed in [All63] to what he called the principle of acquaintance, but was really verificationism; no wonder, since he was a follower of Bergmann, who had been a logical empiricist and never really freed himself from verificationism. Allaire argued in this paper in favour of bare particulars, and assumed that the principle of acquaintance implied that one can only accept the existence of bare particulars if one can show that he is acquainted with them. Allaire completely ignored the possibility that we could know bare particulars by description. Kenneth C. Clatterbaugh in [ Cla65] and Herbert Hochberg in [ Hoc65] and in [ Hoc66] argued against the Principle of Acquaintance misunderstood in such a verificationist sense. The debate
sic terms are derived from experience, from sense-impressions or stimuli via
ostensive definitions, we can yet think and talk meaningfully about entities
of we which we cannot have experience\textsuperscript{38}. All this requires is that we
can combine our basic concepts into complex concepts and judgments and/or
our basic terms into complex terms and sentences in such ways that the
corresponding impressions or stimuli cannot be combined in corresponding
ways in any sensations, and clearly most empiricists have believed this to
be possible. The reason why the sense-impressions or stimuli cannot be
combined in some way can plausibly be explained to in many cases be due
to the nature and constitution of our minds or (as a physicalist would do)
sense-organs or nervous systems, not to the nature of reality as a whole. In
this case the entities our basic concepts are about or our basic terms denote
can be related to entities we cannot in principle have experience of in ways
corresponding to the ways in which the basic concepts are related to the
complex concept and the basic terms are related to the complex term. E.
ge. we can speak about the parts of a minimal observable object a by means
of the complex description ”part of a”, and the minimal observable entity
a can be related to its parts as the term ”a” is related to the term ”part of

centred on the question whether we are acquainted with bare particulars, as they were one
of the most controversial entities claimed to be objects of acquaintance. I will not deal
with the problem of individuation in this work and hence will not discuss bare particulars,
but I will return to the general debate later in 6.4.1. As I will later show, such modern
historians of philosophy as Cheryl Misak also confuse the principle of acquaintance and
verificationism, because they confuse views Russell held at different times.
38We can talk about such entities whether or not we can have propositional knowledge
or justified beliefs concerning them, which is a separate question. Russell already dis-
tinguished (e. g. in [Rus12b, page 72]) knowledge of things and knowledge of truths,
and his knowledge by description is according to him a form of the knowledge of things.
However, Russell yet thought that knowledge by description involved knowledge of truths;
in order to know an entity by description we had to know propositionally that the de-
scription was satisfied by something. We can, however, define a more general concept of
knowledge by description or perhaps better expressed thought by description such that
when we use a definite description we think of the entity that satisfies it, even if we do
not know and even if we cannot know that anything is satisfied by the description. It
is possible to modify verificationism so that statements about unobservable entities are
meaningful (without being reducible to statements concerning observable entities) but no
such statement can be justified. This is basically the view of Bas van Fraassen’s con-
structive empiricism in [vF80]. This view is in my view slightly more plausible than the
verificationism of logical positivists; however, it is also quite implausible at first sight,
and as I will briefly discuss later, van Fraassen’s arguments scarcely suffice in justifying
it in spite of its unintuitiveness. What Ladyman et al. call a non-positivistic form of
verificationism in [LwDSC07, page 29] - not hypothesis which the approximately consen-
sual current scientific picture declares beyond our capacity to investigate should be taken
seriously - is also similar, being a methodological rather than a semantic principle.
Ayer said in [Aye34, page 345] that if a philosopher maintains that his criterion is too narrow and that metaphysical propositions are significant, it is up to him to put forward a more liberal criterion; one that allows the significance of metaphysical propositions yet is not so liberal as to allow for the significance of expressions such as "jealousy pronoun live" or "siffle hip brim". I can surely respond to this challenge by appealing to a view of meaning that seems to follow from the systematic, recursive nature of language, namely the view that a linguistic expression which is grammatical (well-formed) and such that all of the (non-logical) basic terms occurring in it or used to define terms occurring in it have ostensive definitions is always meaningful. This view is still compatible with a weak concept empiricism and far from trivial, for a strong rationalist such as a Platonist (in the exact historical sense) might deny even it, holding that the concept serving as the meaning of some expressions is known to us innately and perceptions only serve to remind us of it. It seems to me that most of what is controversial in verificationism is its rejection of this view, and that logical positivists never gave sufficient reasons to reject this view and hold that even such a grammatical linguistic expression whose basic terms are defined in terms of observables could be meaningless.

This view implies that semantically deviant sentences which might be held to contain category mistakes, such as Chomsky's famous "Colourless green ideas sleep furiously", are meaningful, which many think an unintuitive result. However, it seems to me that the deviance of this sentence is best accounted for by the claim that it is necessarily false and analytically false, but it is still meaningful (and we know that it is analytically false i.e. false because of its meaning because we understand its meaning).

Carl G. Hempel considered in [Hem65, pages 107-113] briefly the possibility of characterizing cognitively significant sentences by demanding that the meaning of all extralogical terms in a sentence be capable of explication by reference to observables. He eventually rejects this possibility and instead went to the other extreme where the criterion of meaningfulness is applied not to terms or sentences but only to entire theories. However, Hempel did not give very good reasons for doing so. He noted that attempts to provide definitions in terms of observables encounter difficulties in the case of disposition terms. However, he certainly did not show that such difficulties would be insuperable. He said in [Hem65, page 109] that a satisfactory definition would require a clarification of the meaning and logic of counterfactual and subjunctive conditionals, which is a thorny problem. However, that it is thorny problem does not show that it cannot be done. In fact, much work towards the solution of this problem has been done by such philosophers as Lewis (e.g. in [Lew73]), Stalnaker, etc, and while difficulties naturally remain, they are surely not greater than those still facing attempts to clarify the verificationist criterion. Unfortunately, much of the debate concerning such conditionals lately has turned to the use of conditionals in natural languages, and the question of disposition terms in natural science has been ignored. However, some philosophers, such as James Fetzer in [FN79],
Part of what is controversial about verificationism is then its commitment to a kind of weak semantic holism, a view according to which it is at the level of sentences instead of terms that the connection between meanings and experience is established. The earlier theories of analytical philosophers such as Russell and probably even Aristotelian theories had been atomistic, at least in comparison to verificationism. This holism, as we will see later in Section 3.4 of this work, was later strengthened to an extreme holism according to which the connection between experience and meaning operates at the level of whole theories, paradoxically when verificationism as a whole was weakened so that verification was replaced by confirmation and disconfirmation.

In fact the consequences of any kind of empiricism, even a far weaker one than verificationism, depend entirely on the theory or description of experience formulated or presupposed by the empiricist. As one special case of this, what problems the verificationist principle declares to be pseudo-problems depends not only on how strong the formulation of the principle itself is, but also on the theory or description of experience formulated or presupposed by the verificationist. It can even be argued that words such as "experience" or "observation" are ambiguous or vague as they are used in natural languages, so that the very meaning of empiricism itself, and a fortiori of the verification principle, is not clear until it is explained what kind of concept of experience is employed in formulating it. If we are unsure just what we can in principle experience or observe, then even if it were correct that a meaningful sentence would have to be in principle verifiable by means of experience or observations, yet knowing this might not suffice
to tell us which sentences are meaningful since we might not know (or even be capable of knowing) which sentences are verifiable in principle.

This description of experience can be called a phenomenology (or more exactly a part of a phenomenology): of course, the positivists would not have liked this word, because they did not like the most famous phenomenology around, Husserl’s phenomenology (as we have already seen in the case of Schlick) and liked even less later existentialist phenomenologies (as we have seen Carnap condemned Heidegger’s writings as nonsense). However, even if their objections to Husserl’s doctrines were valid, the notion of phenomenology should not be tied to the doctrines of any particular philosopher. With the logical positivists the problem of the description of experience took (because of the syntacticism I will examine in the next section of this work) the curious form of the question regarding the form and content of what they called protocol sentences or protocol statements or (more naturally) observation sentences or observation statements basic statements which were supposed to describe the given, or what was observed or sensed. A sentence was supposed by Carnap in [Car32b, page 222] to be verifiable (and hence meaningful) only if could be derived from protocol sentences. It is by no means obvious or uncontroversial how experience, in which verification occurs, should be correctly described; the logical positivists themselves disagreed vehemently among themselves about what the protocol sentences were. The phenomenologists at least tried to approach this problem systematically and rigorously, whether they succeeded or not, but logical positivists such as Ayer did not always even acknowledge the problem. Carnap attempted a systematic description of experience in his Aufbau and Nelson Goodman in [Goo51], so the logical positivists had some systematic theories at the back of their claims of meaninglessness. However, of course these theories could be questioned on many points even by someone who accepted verificationism, so the logical positivists had no right to speak as if the meaninglessness of any particular claim were obvious.

The poorer one thinks experience is, the more sentences verificationism declares meaningless, and conversely, the richer one thinks experience is, the more sentences remain meaningful even according to a strong form of verificationism. The logical positivists had often a very narrow, impoverished conception of experience, limiting it to sense perception. The views of logical positivists that were most flagrantly opposed to common sense
followed usually from such arbitrary limitations on admissible experience. The absurd view that statements concerning the past were meaningless or really spoke about traces in the present followed only with the aid of a view of experience that did not admit memory, even episodic memory, as a form of experience additional to present perception. The logical positivists’ logical behaviourism depended to a great extent on the rejection of inner perception and hence of introspection as a form of observation. Even the classical British empiricists had admitted reflection i.e. the perception of inner mental states as a primitive form of experience; phenomenologists like Husserl had of course a very broad view of experience which included not only such inner perception but the intuition of essences (and as we have seen the debate of Husserl and Schlick turned partly on the correct conception of experience). I will later return in more detail to the question of phenomenology and protocol sentences with regard to the question of whether we have experience of universals in Subsection 6.4.1 of this dissertation.

Also even the way in which the early logical positivists like Ayer (and Schlick and to some extent Carnap - though Carnap’s commitment to the sense-datum theory was a bit more equivocal from the start) conceived sense perception was by no means uncontroversial. In their case the phenomenology presupposed was a form of the sense-datum theory, so that the protocol sentences spoke about sense-data or more accurately sense-contents. However, for earlier sense-datum theorists as e.g. Moore in the papers collected in [Moo22b] or Russell in [Rus12b] the distinction between acts of sensation and sense-data had been central. The early logical positivists denied this distinction so that for them experience did not consist of mental acts directed upon sense-data, but of sense-data or sense-contents themselves. That distinction had of course been the key idea of Moore’s famous attack on idealism and defence of realism in [Moo22c], and therefore rejecting the distinction they naturally also rejected Moore’s realism. However, they did not want to go back to idealism either but instead tried to find a position between realism and idealism, in which I have argued at least some of them (e.g. at least Ayer) failed. So it is not verificationism alone, but only

\[61\] Carnap broadened his conception of experience in [Car56, pages 70,71], admitting that a person’s awareness of his own state of imagining, feeling etc. must be recognized as a kind of observation...
verificationism together with their theory of experience, which led them to the rejection of realism and into neutral monism. However, later many logical positivists, in turning from neutral monism to physicalism following the lead of Otto Neurath took the objects of experience (the entities denoted by terms occurring in protocol sentences) to already be physical objects\textsuperscript{62}\textsuperscript{63}.

This change clearly involved a radical change in the logical positivists’ conception of experience itself and therefore in the very interpretation of the verificationist criterion. Statements concerning sense-data or sense-contents, which earlier were the very paradigm of verifiable and hence meaningful statements, were now often rejected as meaningless and "metaphysical". Nevertheless, even physicalistic verificationism may not be compatible with scientific realism; it allows only statements about macroscopic physical objects like trees and rocks to be meaningful, but not claims about theoretical entities such as atoms or quarks.

It can be argued that trying to combine verificationism with physicalism\textsuperscript{64} may already lead to inconsistency. There are many kinds of reasons for this, of which I will mention three.

On one hand the notion of verification is based on the notion of experience, which Ayer forms an interesting exception as he in \textsuperscript{Aye87} rejected neutral monism and accepted the existence of irreducible physical objects but yet continued to hold that the immediate objects of perception are sense-contents (or rather sense-data, as Ayer returned to the terminology of older analytic philosophy as well as its doctrines).

One prominent motivation for this was that verification should be intersubjective; however, this does not seem to be a sufficient reason for it. While science should be in some way intersubjective, it suffices for such intersubjectivity as is possible that the intermediate rather than the ultimate epistemic foundations of science should be the same for all scientists. Even if the direct objects of experience are physical, yet scientists cannot have the same experiences (or experiences with the very same objects or content) of them. No two scientists can observe the same object at the same time from the very same spatial position relative to it (and indeed usually they observe different though similar objects). Because of this they do not observe the very same facts concerning physical objects and do not get to know immediately the same propositions concerning them. When a scientist replicates an experiment made by another, he replicates the same type of experiment, not the same event of observing, and therefore he has experience of different even if relevantly similar particulars. Another, perhaps more serious reason for rejecting the sense-datum theory lies in claims that the sense-datum theory was wrong even as a phenomenology; I cannot deal with this argument here, and must therefore suspend judgement on the question of the objects of experience.

\textsuperscript{62}It is a bit controversial whether Neurath himself was ever a verificationist, but Carnap (e. g. in \textsuperscript{Car56}) and Hempel at least clearly did try to combine verificationism with physicalism. However, Neurath surely held that metaphysical statements were senseless, and it is hard to see what except verificationism could have been the basis of this view. Therefore if Neurath was not a verificationist, his rejection of metaphysical statements as meaningful probably was wholly baseless.
rience, and it is far from clear how the notion of experience could be understood in physicalistically acceptable terms. The later logical positivists usually tried to do this by appealing to some kind of behaviourism, which was the worst possible way to do it. Neurath already talked about stimuli (Reaktionsprüfung) in [Neu32, page 209] (see [Neu59, page 204] for an English translation). Carnap apparently tried to resolve this apparent inconsistency in [Car32a] (see [Car87, page 458]) by treating men as similar to machines which react to certain situations by displaying signal-disks; this rather crude behaviouristic solution surely alters the meaning of verification very dramatically. Quine (who I will argue was at least at some phases of philosophical development a verificationist) tried to replace talk about experience with talk about stimuli, but there are many problems in this. The words "stimulus" and "stimulation" are ambiguous and can be used in narrower and broader senses. Chomsky showed in [Cho59] that the behaviourist Skinner had fallaciously shifted between the narrower and broader senses of the word. Current verificationists can be accused of doing the same thing (as Chomsky indeed accused Quine of a similar though less gross confusion in the case of the related word "reinforcement" in [Cho69, page 555,56]), though the senses they shift between may be different. It is sure that

---

E. g. in [Qui93, pages 107-108] Quine spoke at first of physical objects and events, then of the impact of molecules and light rays on our sensory receptors i. e. of sensory intake and lastly of neural intake in the brain. Quine was to some extent aware of the danger of equivocation here and tried to avoid it. The discussion in this article, by taking into account the neural input, is far more sophisticated than the crude behaviourism Quine exhibited elsewhere (e. g. in [Qui68]). Indeed, the views Quine expressed in the two articles are hardly even consistent, for in [Qui68, page 186] he denies that a man’s semantics might be determinate beyond what is implicit in his dispositions to overt behaviour, and neural input, which he in [Qui93, pages 107-108] takes as relevant to semantics, is surely determinate beyond what is implicit in overt behaviour. However, while Quine may here have had (at least many of) the pieces for a plausible theory of perception, he did not put them together correctly. The puzzle is more complex than he saw. Quine took in [Qui93, pages 114-115] the sensory intake to be what was essential to perception, instead of neural intake further in the brain (where identity theorists would locate the sensation or perception). Quine gave as the only reason for this that sensory intake is neater, which is a rather weak reason. A more important consideration would seem to be that neural intake in the brain is the only one of the different kinds of stimulus that is necessary for the subject seeming to perceive something, and for his uttering sincerely an observation sentence. Quine seems to have supposed that these three kinds of stimulus are always correlated; he says that the observation sentence is Janus-faced, facing outward to its subject matter and inward to neural intake. However, in such cases as illusion and hallucination, they need not be correlated. The sensory intake can occur without its customary subject matter in the case of illusion and the neural intake can occur without either the customary subject matter or the sensory intake in the case of hallucination. The two faces of Janus can, to extend the metaphor, temporarily separate
the original positivists like Mach or Avenarius would have been horrified by these views, since they were as strongly opposed to physicalism as to traditional idealism.

Adopting physicalism also makes the problem of demarcating between observation statements and other statements far harder. Does observation through telescopes or microscopes count as genuine, direct observation? What about observation through spectacles? Through window-panes?

However, these two problems also face attempts to combine physicalism and direct realism with forms of empiricism weaker than verificationism, so they should perhaps not be stressed too much in the criticism of verificationism. Nevertheless, the behaviourism of logical positivists gave them a much worse chance of solving this problem than more sophisticated physicalistic theories like the identity theory or functionalism would have done.

However the line is drawn, there are strong reasons to think that no statements regarding physical entities are conclusively verifiable (in the way and become different heads which utter different oracles. In such cases the same sensory intake may be associated with different neural input by different speakers. However, it is surely the neural process with which a reductive physicalist such as an identity theorist would identify the process of perception or upon which he would hold the perception to supervene or whose functional role he would hold it to be. Even among logical positivists Herbert Feigl had advocated (in Fei58) such a form of materialism that is surely more sophisticated than the behaviourism to which Quine is committed. Therefore intersubjectivity in practice, which was as important to Quine as it was to Neurath, breaks down. The Argument from Illusion then threatens also the physicalized version of verificationism combined with direct realism. Quine recognizes in [Qui93, page 111] that we are not aware of our neural intake (which I must qualify by saying that we are at least not aware of it as neural intake). In this case it cannot be referred to (as a neural process) in an observation sentence. However, in the case of a hallucination we may not aware of anything objective, but seem to be aware of something because of the neural intake and can sincerely utter an observation sentence. It is easy enough to suppose that in such cases we are aware of a sense-datum; however, there may be other options. One could hold as Place seems to have done at the end of [Pla56] that even in hallucination one seems to perceive the same kind of physical objects that normally cause the neural input, but seems to perceive them in a location where they are not. This kind of causal theory of perception would allow direct realism about physical objects but would also allow observation sentences to be fallible, in accordance with weak foundationalism or foundherentism. Naturally there are many problems also with this view, such as the problem of analysing what “normally” here means in any exact terms and the difficulty of specifying the kind of causal chain that would endow mental states with the object they are commonly thought to be perceptions of. Nevertheless, this view is certainly preferable to behaviourism. I will return to this problem later in Section 3.5.

66Grover Maxwell first attacked the existence of a distinction between observational and theoretical in [Max62], but later accepted it in [Max70]. He could accept it only by assuming that no observational terms denoted physical objects, but the observable is instantiated only in inner events of observers.
that statements regarding sense-contents or sense-data could be argued to be), so the combination of verificationism with physicalism exacerbates the problem of formulating verificationism. It increases the pressure on verificationists to find a formulation that is neither vacuous nor requires conclusive verification, and as we have seen it is far from clear that any such formulation can be found. Unlike the previous problems, this need not be a problem for empiricism or foundationalism as such, since as we have seen a weak foundationalist can take the foundations of knowledge to be fallible, but it is hard problem for the specific strong form of empiricism that verificationism constitutes. Related to this is the problem that taking observable entities to be physical makes drawing a sharp line between observable and theoretical entities or observational and theoretical terms more difficult; sense-data would be so much different from physical objects that few predicates could apply to both, so drawing a line is quite easy in their case. Also the coherentism that Neurath has been interpreted as holding makes the problems still more difficult, for coherentism seems inconsistent with both verificationism (since it is inconsistent with foundationalism and verificationism is a version of empiricism which is a version of foundationalism) and physicalism (since theories which claim that non-physical entities exist or even that physical entities do not exist can surely be coherent and hence according to this theory they would be true).

As this debate continued, many logical positivists came to think that the choice of protocol sentences was a matter of convention, as seen for example in [Car32a] (translated into English in [Car87]) and even more extremely in [Neu34, page 348] and in [Hem00, page 18]. However, this view threatens to amount to a complete rejection of empiricism, even of moderate empiricism, and of all kinds of foundationalism, ending in an absurd epistemic relativism where "anything goes". If protocol sentences can be chosen arbitrarily by convention, as Neurath and Hempel say, then why not choose sentences from the Bible or Quran as protocol sentences, as religious fundamentalists would like to do? Why not take as a protocol sentence that Noah saw that water covered all the Earth? Neurath’s coherentist and conventionalist version of logical positivism is then an utter failure as a defence of scientific philosophy; it allows young earth creationism to be just as well justified for creationists and even "true" as evolutionary theory is for evolutionary theorists.
Because of this a defender of a scientific world-view must accept foundationalism rather than coherentism or conventionalism, at least if these exhaust the options. However, even if these are the only options, the foundationalism he has to accept need not be a strong foundationalism such as Schlick’s, but can be weak foundationalism which is compatible with fallibilism. Susan Haack has suggested a third option, which she calls foundherentism. However, it has been argued, at least by Lawrence BonJour in [Bon97] and by Peter Ramel in [Tra08] that Haack’s foundherentism is really a version of foundationalism, what Haack and Tramel call feeble foundationalism. BonJour agreed that Haack had shown coherentism, which he had supported earlier, to be untenable, but instead of converting to foundherentism he converted to foundationalism. The question is whether foundationalism can allow that basic beliefs are (further) justified by other basic beliefs and even by derived beliefs. This may be just a verbal question; it does not much matter what an epistemological position is labelled. The important point is that we can accept some of the important features of traditional foundationalism and combine them with features of traditional coherentism67.

It is indeed possible that at least some of the logical empiricists speaking out on behalf of conventionalism did not intend that the choice of protocol sentences would be completely conventional, and of course a moderate conventionalism would be defensible, if it could be restrained within reasonable limits. Obviously it is possible to develop syntactically different languages in which protocol statements can be formulated, if they are semantically equivalent in some strong sense. However, it seems quite plain that Neurath and the early Hempel at least did not avoid complete relativism68. I will later return briefly to these fundamental epistemological questions in con-

67Bertrand Russell seems to have anticipated feeble foundationalism or even foundherentism, as he said in [Rus40, page 150] that there may be evidence in favour of a basic proposition, but it is not this evidence alone that causes the belief in such a proposition.

68Carnap accepts in [Car63b, page 864] that some of Neurath’s formulations are misleading, but thinks that Neurath never really held extreme conventionalism or the coherence conception of truth. It must, however, be suspected that Carnap’s interpretation of the thinking of his fellow positivists was too charitable. Carnap stresses that his Principle of Tolerance did not refer to the content of synthetic sentences. However, this cannot be correct; if the inference rules could be chosen arbitrarily, as Carnap claimed in [Car37, page xv], then we could also choose to hold any synthetic sentence p true just by adopting an inference rule according to which p can be inferred from any sentence at all. Besides even if this had been true of all of Carnap’s theories, this would not be enough, for Quine’s argument in [Qui76] shows that it is wrong even so far as it refers to the content of logical theories.
nection with the problem of universals in 6.4.1, though it is not possible in this work to attempt any kind of comprehensive solution of them.

3.4 Holistic Verificationism and Semantic Holism

While the stronger version of verificationism was actually a theory of what meaning was, namely a method of verification, it is not clear whether the weaker form of verificationism could yield such a theory (even if it turned out not to be completely vacuous, which is, as we have seen, doubtful). This is an important question since many modern verificationists say that meaning could be understood in terms of verification conditions or assertibility conditions (as I will later show Hartry Field does) or acceptance conditions etc. (there are many expressions for the same idea), and one must ask if this commits them to a strong form of verificationism or not. A natural way to generalize the idea of meaning as a method of verification would be to hold that meaning is a method of defeasible confirmation or disconfirmation. However, this appears to lead to some very strange and unpleasant conclusions. It has often been argued (as was already done by Pierre Duhem) that confirmation is holistic in nature, and if that is correct then this generalization would lead to a view according to which meaning is also similarly holistic, a view called (strong) meaning holism or semantic holism. According to this view individual sentences are not meaningful in themselves, but only complete scientific theories are.

Semantic holism was professed by many philosophers within the logical positivist movement or influenced by it, e.g. by Carl Hempel (e.g. in [Hem65, page 113]) and as I will argue by Quine. However, it leads to many very unattractive, indeed disastrous conclusions. It makes it very hard to compare theories, since according to it the same term - or at least the same non-observational term - occurring in differing theories always has a differing meaning in them, so it leads to well-known problems of incommensurability. It even makes it mysterious how people with different beliefs

\footnote{One consequence of verificationism and holism drawn out by Hempel is the view according to which scientific theories are partially interpreted calculi, a view also defended by Carnap in [Car56] and by other philosophers such as Ernest Nagel and Braithwaite. Hempel considered in [Hem65, pages 114-116] several suggestions that consider theoretical systems to be cognitively significant if they are partially interpreted to some extent by means of what are called bridge laws or correspondence rules, and while he did not ultimately accept any of them, yet this kind of suggestion has become widespread. Indeed it...}
and thus different theories are at all able to communicate. Nevertheless, it has been argued by Peter Achinstein in \cite{Ach63} that it presents no non-trivial restrictions on what theories are meaningful, and allows metaphysical theories to be meaningful along with any theories at all.

It may indeed yet allow us to compare theories in some (unclear and probably very cumbersome) way as wholes by comparing their observational consequences as wholes, so the incommensurability this kind of semantical holism implies by itself is not quite as bad as those implied by some other theories in the philosophy of science; nevertheless, it still seems quite unacceptable to me. However, if this kind of semantical holism is combined (as it has often been combined) with the popular view that it is impossible to draw any line between observation language and theoretical language, and that observation is theory-laden (as Thomas S. Kuhn argues in chapter X of \cite{Kuh70}), this leads to full incommensurability of the kind Thomas S. Kuhn famously argues for in \cite{Kuh70}, a situation where any kind of comparison between theories is impossible.

It must also be noted that this kind of semantic holism is very odd as an outcome of analytical philosophy, since it would make any kind of analysis impossible. Indeed, it resembles very much the views of the monis-

---

has become known (as e. g. in \cite{Sup00}) as the Received View or the Standard Conception of scientific theories. It is also sometimes called the Syntactic Approach or the syntactic view, though what is distinctive of it is not its taking theories to be syntactic entities, but the way in which theories are interpreted according to it (and it is clearly distinct from the view according to which theories would be sets of wholly uninterpreted sentences, which would most appropriately be called the syntactic view). This Received View is sometimes described simply as the view of logical positivists, though it was a view characteristic only of the last stages of logical positivism; this may lead to a misunderstanding of logical positivism as a whole, since most views of theories developed within it were based on a less holistic kind of verificationism. However, sometimes definitions of theoretical terms by means of observational ones are considered as limiting cases of correspondence rules, and in this expanded sense the Received View might cover a lot of logical positivism. The obvious problem with the view is that intuitively it is clear that no theoretical system which is only partially interpreted is significant (at least in an unequivocal sense; at most such systems are partially significant). The Received View is often called the Syntactic View and contrasted with what is called the Semantic View, in which theories are taken to be sets of models. However, it is by no means clear that the "Semantic View" is the only or the best alternative to the Received View. As Anjan Chakravartty argues in \cite{Cha01}, if independence from language is the desideratum, why bother with models; wouldn’t propositions suffice? The Semantic View also seems to have holistic and structuralist features, so its difference from the Syntactic View may not repair what can be argued to be the worst defect of the Received View. At least a theory cannot be taken to be the set of all of the models of its formulations, for this would amount to taking it as an abstract structure, and this would amount to the kind of structuralism that as we saw Newman has shown to be untenable.

---

138
tic Hegelian idealists as a protest against whom (by Russell and Moore) analytical philosophy was born. In arriving at semantic holism analytical philosophy has then in a way nullified itself. Something has then gone terribly wrong, and it is surely plausible that the first and most important thing that went wrong was the acceptance of verificationism itself.

Among philosophers who have opposed semantic holism have been Jerry Fodor and Ernest Lepore in [FL91], Michael Devitt in [Dev93] and others. Fodor and Lepore defend a rather extreme semantic atomism, while Devitt argues for a prima facie more plausible, moderate semantic localism. I find the criticism of Devitt more to the point than that of Fodor and Lepore (though I cannot delve into the detailed arguments here), but they show together that there are many points from which the arguments for semantic holism can be attacked. Even some verificationists such as Michael Dummett have been appalled at the consequences of implementing the verificationist program in this way (though it is far from clear how Dummett could escape from holistic consequences himself). These philosophers often disagree on what is wrong in arguments for holism but most see the disastrous consequences of holism as sufficient to doubt the premises of any argument for it.

Other arguments than the one based on verificationism have also been given for semantic holism - e. g. holism has been defended on the basis of (crude versions of) functionalism. However, it seems to me that the argument based on verificationism is by far the most important of the arguments for semantic holism, so if verificationism is rejected then the case for semantic holism becomes significantly weaker. One common and important argument for semantic holism that is often held to differ from the argument based on verificationism is based on Quine’s criticism of the distinction between analytical and synthetic statements. However, Quine’s argument

\[70\] Though the absence of the distinction is commonly used as an argument for semantic holism, Fodor and Lepore strangely purport to use the absence of the distinction as a premise in an argument against semantic holism. However, it seems to me that the crux of their argument against semantic holism lies in their appeal to compositionality, and what the absence of the distinction really does is motivate them to argue that if semantic holism is false, then one must accept their extreme semantic atomism. If my interpretation that Quine’s argument for the absence of the distinction depends essentially on verificationism is correct, then one who rejects verificationism entirely need not accept either semantic holism or Fodor’s and Lepore’s extreme semantic atomism. He can reject both, which should be a welcome result, since both are intuitively extremely implausible positions with unwelcome consequences.
is far from being completely independent of verificationism. This may be somewhat controversial, as there have been attempts to argue that Quine was not at all a verificationist, as Panu Raatikainen’s argument in [Raa03], while other interpreters of Quine such as Roger Gibson view him as definitely a verificationist, so I will have to defend this view at some length.

As Raatikainen says in [Raa03], naturally the whole issue of whether Quine should be considered a verificationist depends essentially on one’s understanding of the term ”verificationism”. Unfortunately, as we have already seen, a very great number of formulations of the verificationist criterion have been proposed, so it is by no means a simple task to say what is essential to verificationism, but involves a difficult task of explication in Carnap’s sense. If we stuck to the very earliest formulations of verificationism, then the later view of logical positivists that the meaningfulness of sentences depended on defeasible confirmation rather than conclusive verification would not count as verificationist, but it is clearly usually counted as verificationist, and this seems quite correct to me. Therefore we must be willing to generalize the formulation somewhat, and it is not clear where the line appears where we just start talking about a different matter. There is also the possibility that Quine’s views with regard to verificationism may have changed significantly between different writings, perhaps even in ways Quine himself refused to acknowledge or did not even realize himself. Therefore instead of asking whether Quine was a verificationist we should ask just whether he was a verificationist at a certain time or in a certain article or book. Thirdly, Quine’s views may have been at some time so radically inconsistent that it cannot be said whether he was a verificationist then. If we look at all the quotations assembled in [Raa03], both those taken from Gibson in favour of Quine being a verificationist and those gathered by Raatikainen himself in opposition to it, it is far from clear that the views expressed in them can be combined consistently. There are then good reasons to suspect that their diversity reflects either a change of views or permanent inconsistency on Quine’s part. Because of these three reasons I agree with Raatikainen that one cannot straightforwardly or unqualifiedly take Quine to be verificationist. However, I will argue that it is less misleading to count the views Quine expresses in [Qui53e] and many other writings as verificationist than to deny that they are verificationist, though I will leave open whether the same holds of other writings of Quine.
If we look carefully at the argument in his famous article [Qui53e], we will see that the argument essentially depends upon verificationism in a crucial point, so Quine surely was a verificationist in a broad sense at least when he wrote that article. In the beginning of the article Quine attacked the distinction by arguing that various attempts to analyse it in more primitive terms are unsuccessful. However, this kind of argument surely does not suffice to show that any distinction would be bogus, at least if the distinction has some presumption in its favour (as e. g. Grice and Strawson already argued in [GS56, page 142]). If Quine’s argument as a whole is to be taken seriously, then this initial argumentation can only be a preliminary to the real nerve of Quine’s argument, which depends on verificationism.

Quine’s final attempt to define analyticity and synonymy in [Qui53e, §5, page 37] appeals to the verification theory of meaning (in its older phenomenalistic or neutral monistic form). Quine clearly saved for the last the attempt which he found to be the most plausible. Quine explained that if the verification theory were correct, then two sentences would be synonymous if and only if they are alike in point of method of empirical confirmation of information. However, Quine rejected this attempt because of problems in Carnap’s old atomistic and phenomenalistic form of verificationism in the Aufbau (though the holistic verificationism Quine arrives at in this article seems to me to resemble greatly Carnap’s later view in [Car56]). It is only because of verificationism that Quine ties the two ”dogmas” together. Quine showed that if verificationism is correct (as he assumes) and the distinction between analytic and synthetic sentences is genuine, this implies that reductionism is correct. However, Quine argues that reductionism is false, on the rather weak ground the best attempt to implement it, Carnap’s attempt in the Aufbau, has not succeeded in carrying it through. Therefore, by modus tollens, either verificationism is incorrect or the distinction is not genuine. Quine thought that the first option was unacceptable and hence deduced that the distinction was not genuine. However, if we reject verificationism, we can consistently both accept the distinction between analytic and synthetic statements and reject reductionism.

However, it seems to me that Quine did not reject the attempt to illuminate the notion of analyticity with the aid of a verification principle, but only presented as his solution (as we have seen Hempel also did) that the verification principle must be stated in a different, holistic or organic form,
and his revision of the concepts of analyticity and syntheticity follows from this. Hempel had already formulated the verification principle in a holistic form, but had yet clung to the distinction between analytic and synthetic statements, as can be seen in the very first page of [Hem65, 101].

I would like to propose that what is essential to verificationism (or at any rate objectionable in it) is the view that any acceptable semantic notions must be defined with the aid of epistemic notions like verification, confirmation, etc. If we defined verificationism strictly by referring to meanings of sentences, then it would be hard for two reasons to apply the definition to Quine’s views, whether to accept or deny that he is a verificationist. First of all as Raatikainen notes Quine repudiates the very notion of meaningful, if understood as having a meaning. Secondly since Quine denied that sentences in general - Quine excepts observation sentences from this claim at least in some of the phases of his philosophy - have independent meanings, he could not say that the meaning of sentences was their method of verification of confirmation (other than in the case of observation sentences, in whose case even an opponent of verificationism might admit that claim). However, both reasons can be circumvented. To the first point it must be noted Quine yet seems to retain some semantic notions as acceptable, and truth is one of them. Therefore my proposal is a natural generalization of the customary definitions that is applicable Quine’s framework.

As to the second reason, even if Quine could not say that the meaning of a sentence is its method of verification or confirmation, he could yet say that the meaning or significance of something is its method of verification or confirmation. The distinction between verificationist and non-verificationist theories can also be made in the case of holistic theories of semantics. Since Quine thought that the unit of significance was not the sentence but the theory, a cluster of sentences, we must apply the definition of verificationism to theories. We must ask whether he thought that verifiability or confirmability was for him the criterion of the meaningfulness of theories and whether the meaning or significance of theories is their method of verification or confirmation. The answer seems to be that he did, and that suffices to make his theory verificationist. Quine would surely have said that such a theory that neither it nor any consistent extension of it implies observation categoricals (or help to integrate any lower-level theory that does imply them) is meaningless, while a theory that implies them is meaningful or significant. An
anti-verificationist would, however, say that even a theory which does not imply observation categoricals is meaningful if the sentences contained in it are grammatical and all primitive terms occurring in sentences contained in it or in definitions of its terms have an ostensive definition (i.e. occur in observation sentences).

Let us see what happens when we apply it to Quine’s views in the article we are examining. Quine says in [Qui53e, page 41] the following regarding the truth of statements:

The factual component must, if we are empiricists, boil down to a range of confirmatory experiences.

The kind of empiricism Quine exhibits here definitely counts as a species of verificationism in the sense in which I proposed to use the notion, since truth is a semantic notion and Quine here connected to it the epistemic notion of confirmatory experiences. This view is unacceptable to a scientific realist or to anyone who is a correspondence theorist about truth, since they must hold that the factual component contains unobservable entities (i.e. entities that cannot be contents or objects of experience) like atoms or electrons. I suppose it would objectionable to Raatikainen himself (who does not refer to this passage and may not have noticed it), as he has on many occasions defended scientific realism. It is surely objectionable to Michael Devitt and Panu Raatikainen, who are scientific realists and hold a correspondence theory of truth, so I cannot see how they can consistently appeal to Quine in defending an extreme holistic empiricism.

71 Devitt claims that one can consistently hold together with Quine that all beliefs are epistemically revisable without accepting Quine’s semantical holism. Such a combination of semantic localism and epistemic holism might be consistent, as its inconsistency has not (so far as I know) been proved. However, this does not make it certain that such a view is consistent; one would require some sort of independent argument for inconsistency, which has not (again so far as I know) yet been provided either. Nevertheless, even if such a view is consistent it seems to me that Quine’s arguments for the revisability thesis are not separable from his arguments for semantical holism, so Devitt cannot consistently appeal to Quine’s arguments for the revisability thesis while rejecting semantical holism. Confirmational holism as already argued for by Duhem does not amount to the full revisability thesis, for even if confirmation is epistemically holistic, it does not follow that all statements are epistemically holistic, for analytical statements would usually not be taken to require confirmation. I do not think that Quine has any arguments for the claim that they require confirmation that would be separable from his arguments for semantical holism.
3.5 Current Verificationism and Logical Positivism

We have seen that there are very good reasons to think that the verificationism of logical positivists has been shown to be quite unsustainable. Yet almost no one draws from this the obvious conclusion, namely that if the main reason the positivists had for opposing metaphysics was wrong then probably the positivists had no rational cause to think metaphysics illegitimate, so metaphysics, even transcendent or speculative metaphysics, may well be a quite legitimate area of study.

This is perhaps because different objections to logical positivism, objections proceeding from incompatible premises, have been confused with each other. There are two different ways in which later philosophy has diverged from logical positivism, which offer incompatible objections to logical positivism. On the other hand there have been scientific realists who have objected to the positivists’ rejection of realism, and on the other hand there have been various kinds of relativists, who have diverged even farther from realism than logical positivists ever did. It is characteristic of the confusion in which the temporary demise of logical positivism left analytical philosophy that many analytical philosophers have at different times rejected logical positivism as not realistic enough and as too realistic.

3.5.1 Quine and Kuhn and Logical positivism

Quine and Kuhn are commonly thought to have vanquished logical positivism in [Qui53e] and in [Kuh70], as e. g. Friedman says in [Fri91, page 505]. However, it seems to me that they by no means completely overcame logical positivism, at least in those writings, but rather transformed it in many respects into a more extreme form in them.

I have argued that the trouble with logical positivism is that it was too strongly empiristic and too holistic; however, Quine replaced it with a theory that was even more radically empiristic and holistic. This may have helped to remove some small problems and inconsistencies, but the resulting theory has the same major defects as the philosophy of earlier logical positivists.

Friedman says in [Fri91, page 505] that philosophers of science have characteristically proceeded on the basis of Kuhn’s well-known critique. This implies that scientific realism, which cannot agree with Kuhn’s critique, would be wholly uncharacteristic of the philosophy of science. However,
Friedman scarcely provides any evidence for so dramatic a claim. Surely there have been many influential scientific realists lately; in this work alone I mention Niiniluoto, Chakravartty, Devitt, Ellis etc. Indeed, I doubt whether we could say that any single procedure has been characteristic of philosophy of science after the collapse of positivism, as it has fragmented into several small schools, and most philosophers of either science or language seem to vacillate between more and less realistic theories.

So far as it is scientific realism that is thought to replace logical positivism as the interpretation of science even among analytical philosophers, then the prospects for not only metaphysics in the broadest sense, but even for speculative metaphysics, would seem to be bright. One would have expected that in rejecting the logical positivistic interpretation of science scientific realists would have also rejected the positivists’ opposition to (even transcendent) metaphysics, since the latter was based on the former. This sometimes happened, as seen in the very title of some works, such as Anjan Chakravartty’s [Cha07] A Metaphysics for Scientific Realism or Brian Ellis’s [Ell09] The Metaphysics of Scientific Realism. However, scientific realists are not always willing to go along with this since they have often retained the positivists’ hostile attitude to metaphysics even when all justification for it had passed. However, if it were relativism of Kuhn’s brand that replaced logical positivism (or arguably rather consummated it), then there might be rational reasons to despair of the possibility of at least speculative metaphysics.

When Passmore said in [Pas67] that logical positivism was dead, he did not mention Quine as having defeated logical positivism, but only as a visitor to the Vienna Circle (though the article [Qui53e] had by then appeared many years earlier). Nor did Passmore mention Kuhn, though Kuhn’s major work [Kuh70] had also appeared before the time he wrote. This suggests that Passmore did not think that the defects in logical positivism were those Quine or Kuhn attacked; Passmore’s autopsy of logical positivism delivered a different cause of death than Friedman’s. One of the doctrines of logical positivism that Passmore thought was dead was verificationism. Even if Quine were right that the distinction between analytic and synthetic is problematic, surely verificationism is even more problematic, for well-known reasons many of which we have gone through. While the distinction between analytic and synthetic has a presumption in its favour, verificationism does not
have one, as it is explicitly revisionary. Schlick did indeed try to argue that verificationism would have a presumption in its favour, but I have already argued that he was mistaken. Also the main problems with verificationism (as well as the main problems with structuralism) are quite independent of the purported problems with analyticity. On the other hand verificationism is distinctive of logical positivism, and I have argued that Quine did not reject verificationism. The unity of science is another doctrine that is commonly held to be distinctive of logical positivism and Quine did not reject that either. Structuralism is yet another common logical positivist doctrine that Quine held, as seen in such articles as [Qui08a]. Therefore Quine cannot be said to have ever rejected logical positivism completely but only to have created a new form of it. Many philosophers who say that logical positivism is dead and think that Quine killed logical positivism are really themselves logical positivists. Contrary to what e. g. Friedman thought, logical positivism is alive and well today - unfortunately, as its doctrines have been either definitely disproved or disconfirmed several times. Because of this I have not just been flogging a dead horse in presenting in detail the arguments against verificationism in this section of my work.

Because Quine argued in [Qui53e] against reductionism, more specifically the kind of reductionism according to which physical objects can be reduced to immediate experience, a view that scientific realists are also unhappy with, the view he defended there might seem similar to the view of scientific realists that physical objects exist independently of human experience. This impression is naturally reinforced by the fact that in other, later writings Quine does seem to rather unequivocally defend views characteristic of scientific realism. However, it would be a mistake to understand Quine’s influential article in this way; the reasons Quine had for objecting

---

72Opposition to metaphysics might also be argued to be distinctive of logical positivism, and Quine did relax opposition to metaphysics, so in this one respect Quine did overcome logical positivism. However, this is just one respect compared to many in which Quine remained confined by the dogmas of logical positivism.

73There is, indeed, some obscurity in Quine's relation to structuralism. Quine declares himself in [Qui08a, page 405] in favour of a global ontological structuralism; however, he then says in [Qui08a, page 406] that his global structuralism should not be seen as a structural ontology. This is either totally incoherent or utterly obscure; what possible difference could there be between ontological structuralism and structural ontology? Quine proposed this distinction in order to escape an incompatibility between structuralism and naturalism he himself saw to threaten his theory; however, if we cannot make consistent sense of the distinction, then we must conclude that Quine did not succeed in reconciling structuralism with naturalism.
to reductionism in [Qui53e] are exactly the opposites of those a scientific realist would have. So are the reasons Kuhn has for objecting to logical positivism; Kuhn carried further the logical positivist objection to the atomism and realism of original analytical philosophers.

While Quine rejected the atomistic form of Carnap’s verificationism, he did not in this article reject Carnap’s phenomenalism, though Carnap himself and most other logical positivists had long ago (under the influence of Neurath) rejected this. A the end of [Qui53e, page 19] Quine had left the choice between phenomenalism and physicalism an open question. In [Qui53e] he seems to come decisively in favour of phenomenalism, though in a novel way. Rather than claiming that physical objects exist independently of experience, Quine claimed that physical objects are not even logical constructions out of sense data. A reductionist theory that explains physical objects as logical constructions can yet admit that positive statements concerning physical objects are literally true (when properly interpreted). Quine’s theory, however, appears to have been instrumentalist; he said in [Qui53e, page 44] that he thought of the conceptual scheme of science as a tool for predicting future experience in the light of past experience. An instrumentalist about physical objects cannot allow that statements concerning physical objects would be literally true. Quine claimed in [Qui53e, page 44] that physical objects are nothing but myths, comparable to Homer’s gods. Quine’s holism led him to a very radical view (see [Qui53e, §6, page 45]); when Carnap had maintained that an ontological question such as the question of the existence of numbers was not a question of matters of fact, Quine apparently agreed but said that in that case the same holds regarding scientific hypotheses generally. If Quine really agreed with Carnap and extended his view to all scientific hypotheses - and not just agreed provisionally for the sake of argument, which is also a possible reading - then his view is not compatible with scientific realism, much less with naturalism or physicalism (which serves to further confirm my suspicions that verificationism and physicalism are incompatible). Could any physicalist be happy with the claim that physical objects are myths? A scientist (whether physicalist or not) would surely be even less happy to hear that such questions as whether electrons or even planets exist are not questions of matters of fact.\footnote{The comparison of physical objects to Homer’s gods is ambiguous, and since Quine here speaks of epistemological footing, he could be interpreted (and has been interpreted...}
People can be misled by the fact that just rejecting the distinction between analytical and synthetic statements or accepting extreme empiricism regarding justification might not by itself lead to this kind of extreme anti-realism and instrumentalism (as e.g. Devitt’s defense of extreme empiricism regarding justification may not lead him to such an anti-realism). However, if the distinction is rejected for the reasons and by the way that Quine rejected it in [Qui53e], namely because of and through a holistic and verificationistic theory of meaning, then this does lead to anti-realism and conventionalism. Though Quine’s philosophy, like that of the logical positivists, started out by professing great respect for science, it was yet in the end as dismissive of the true objective significance of scientific theories as the logical positivists turned out to be.

Of course, Quine later changed his view in this respect, accepting physicalism for example in [Qui93] and speaking of stimuli rather than sense-data. However, this might not constitute a complete abandonment of logical positivism, any more than his earlier turn to holism, but just a change from the kind of positivism earlier espoused by Schlick, Ayer and the early Carnap to the kind of logical positivism that had earlier been espoused by Otto Neurath and Carl Hempel and the late Carnap. Kuhn also seems to represent a later version of Neurath’s kind of logical positivism since like Neurath he opposes both foundationalism and the correspondence theory of truth in [Kuh00, page 95]. Both Kuhn and the late Quine then seem to represent

by charitable commentators) as just saying that both are inferred entities, not as denying that one class of entities exists and the other does not. However, in that case the use of words like "myth" is surely rather misleading, since the word "myth" in its ordinary sense surely means (when applied to a story) something that is not true in a literal sense, though it may be true in a metaphorical sense. Originally the word "myth" in Greek meant a story concerning gods, and such a story could be literally true. However, as with the rise of philosophy traditional myths concerning Olympian gods were increasingly taken to be false in their literal sense, and Greek theologists tried to defend them by claiming that they were intended to be taken metaphorically (a defence later adopted also by liberal theologists of other religions), the word acquired its currently ordinary sense. When applied to an entity the word "myth" therefore means something that does not exist in the literal sense. Quine’s use of the word "posit" has been concentrated on in discussions of Quine’s realism or anti-realism, but this attention to it seems overweening; that word can surely be understood in a sense compatible with realism. However, the use of such words as "myth" is far worse. Feyerabend also compared scientific theories to myths, so Quine here prepared the way for Feyerabend’s relativism. In any case, saying that the truth of a scientific hypothesis is not a question of fact is a quite unambiguous rejection of realism concerning scientific hypotheses. It cannot be understood as a mere epistemic claim but is clearly an ontological claim. Presumably questions concerning sense data would have been questions of matter of fact in this stage of Quine’s philosophy.
and carry further some of the variant tendencies of logical positivism, rather than fully overcoming logical positivism.

However, there is one problem here in the interpretation of Quine. It has been debated whether Neurath was or could consistently have been a realist concerning theoretical physical entities such as quarks or electrons, or only of observable physical entities such as rivers and trees\textsuperscript{75}. The same question can of course be raised about Quine. However, Quine did rather clearly commit himself to realism about theoretical entities in some later writings; e. g. in [Qui08a, page 406] he said that his tentative ontology continues to consist of quarks and their compounds. However, it can be asked if this is consistent with his semantical and epistemological principles, if he continued to hold verificationism and structuralism. It is not at all clear how one could give stimulus meaning for a theory speaking about quarks, if the word "stimulus" is used in any sense of the word narrow enough to be interesting (e. g. if quarks themselves could be held to be stimuli, then clearly the meaning of the word "stimulus" would have become utterly vacuous and have nothing in common with any concept usable in psychology). Can Quine mean by the claim that quarks exist any more than that the statement that they do can be used to derive true observation categoricals? Surely when scientific realists (and philosophically uneducated scientists) claim that there are quarks, they mean far more than that. However, if Quine did not continue to hold verificationism, then it can be wondered how good reasons he could have had to continue to hold semantic holism - which it is in any case clear he continued to hold - since as I have shown he had originally justified semantic holism mainly with the aid of verificationism. At the very least his case for semantic holism would have been weakened significantly with an abandonment of verificationism even if not completely evaporated. Quine’s structuralism also seems to be incompatible with his postulation of quarks since it follows from Newman’s argument that quarks cannot be distinguished from other entities structurally, as a proxy function can transform a set of quarks to any set with the same cardinality. There are then reasons to suspect an inconsistency even in Quine’s later philosophy between positivistic and realistic elements it combined.

\textsuperscript{75}If he was not a realist concerning such entities as quarks which physical theories speak about, then his view should not really be counted as physicalistic in the most common sense of the word, but only as materialistic.
Interestingly, later still Quine in several places (as in [Qui08b, page 393] and in [Qui93, page 111]) seemed to back down from an extreme holism and even opposed the incommensurability thesis of Kuhn and others. He claimed no that the unit of significance was not the whole of science, but clusters of sentences big enough to imply observation categoricals. However, some of Quine’s claims still seem to imply an extreme holism, and one can doubt whether it can really be distinguished from Kuhn’s theory as Quine himself supposed. Quine held in [Qui93, page 109] that observation sentences can function as observation sentences only holophrastically, so that the terms occurring in them have no definite denotations. However, applying holism to observational statements is not justified even by verificationism. Furthermore, the consequences of such a move may be disastrous in several different ways. We can adapt a point made by Chomsky against Quine in [Cho69, page 57]; this would imply that observation sentences should be learned one by one (if they were to be used genuinely as observation sentences) and therefore does not appear to be consistent with the obvious fact that a person learns to use an infinite number of observation sentences (without them becoming theoretical sentences). It is even worse that Quine takes it in [Qui93, page 112] to license the application of proxy functions to their content, and this leads to the kind of structuralism that as I have related Newman showed to be untenable.\footnote{It is not wholly clear that this move is valid even if observation sentences were holophrastic; a proxy function or isomorphism need not preserve even an undifferentiated range of neural intake.} Quine’s reason for taking observation sentences to be holophrastic is that if the terms occurring in an observation sentence had definite denotations, the sentence would then in his view be theory-laden (as Kuhn held); however, we can doubt whether this would have to be the case.\footnote{One odd feature in Quine’s theory, inherited from Neurath, is that he held observation sentences to be learned in early childhood (and therefore stated in an ordinary language). This would indeed make them theory-laden. However, this is an odd view for Quine to take, since Quine elsewhere follows an artificial language method rather than an ordinary language method in philosophy. The role observation sentences are supposed to play in the theory of science is that they are sentences in which the results of scientific experiments are formulated. Surely the language in which such results are formulated ought to be more exact than the language which a child first learns. In fact, in an ideal case it ought to be an artificial language. It is indeed likely that as Neurath argued a completely precise ideal language is impossible to construct, but this does not imply that it would be impossible to construct artificial languages far more precise than any ordinary or natural languages, and in fact the languages of exact sciences like physics are already to a great extent such. It has not been shown that the sentences of such an artificial language must be theory-laden;
We could in any case avoid Quine’s disastrous conclusion (even if we were physicalists) by hypothesizing that even when the observation sentence faces inward different terms occurring in it could be associated with different parts of the neural intake\(^{78}\). Of course, only the physiology and psychology of a far future could definitely show if such a hypothesis is true, but the same holds of Quine’s original claims about the role of stimuli.

### 3.5.2 Dummett’s Justificationism

Logical positivists thought that irrational metaphysical theories threatened natural science and thought that they could use verificationism to defend science. Verificationism, however, seems to have been originally inspired by the thinking of Wittgenstein, who was not at all as enthusiastic about science as most of the logical positivists. When logical positivists saw that verificationism threatened science as well as metaphysics they were willing to practically abandon verificationism, even if they did not usually openly admit that they had wholly reject their previous views. However, later verificationists have been willing to bite the bitter pill and sacrifice huge amounts of science and mathematics to the altar of verificationism, being in this more in tune with Wittgenstein’s mode of thinking.

For example, the philosophy of Michael Dummett, as developed for instance in [Dum91], has been very influential, though Dummett, following the intuitionists, has to sacrifice not just a lot of theoretical physics, but even much of generally accepted mathematics and logic. The early logical positivists generally thought that mathematical propositions were analytic and hence verificationism was irrelevant to mathematics, as it only concerned synthetic propositions. They mainly followed logicists in their philosophy of mathematics, though with a more conventionalist gloss. However, some of the earlier thinkers who had influenced logical positivists such as Poincare had opposed logicism and instead anticipated intuitionism. Wittgenstein’s

---

\(^{78}\)Even in the case of hallucination, when the terms occurring in an observation sentence are not associated with any external entities, they could be associated with definite parts of the neural intake, even though the observer is not aware of this. This could be argued to give them a definite reference; they could refer to the entities by which the neural intake associated with them was originally or is typically caused, even when they are caused by different causes (which is close to the theory that Place suggested in [Pla56]). They would then denote physical entities without yet being intersubjective in practice.
original views had according to many commentators been influenced by intuionists (see e.g. [Rod11]) though most logical positivists when influenced by them had disregarded this aspect of them. Dummett now brought these intuitionistic ideas into verificationism and thought that verificationism demanded radical changes even to mathematical practice. Dummett called his theory justificationistic rather than verificationistic, but it is clearly in the verificationistic tradition.

Dummett did not indeed use the principle to attack metaphysics so much as to defend (what he describes as) a specific metaphysical theory, metaphysical anti-realism: he is more perspicacious or more honest in this respect than the logical positivists, since as we saw their opposition to metaphysics concealed an implicit anti-realistic metaphysics. However, unlike Ayer’s implicit metaphysics, Dummett’s metaphysics is not a version of neutral monism but a far more radical sort of anti-realism. Indeed, Dummett went so far as to use in [Dum06, chapter 8] his anti-realism as a premise in an argument for the existence of God just like Berkeley, though unlike Berkeley Dummett did not definitely commit himself to the conclusion of this argument. Dummett argued in [Dum06, page 96] that the concept of the world as a whole is correlative to that of God, as standing over against the world and if this contrast is removed, no room remains for distinguishing the world as it is in itself from the world as we experience and find it. This is rather ironic as such verificationists as Ayer had claimed that sentences apparently speaking about God were meaningless, and it is far from clear how people could manifest their understanding of sentences speaking about God. Clearly Dummett was not trying to use verificationism to defend a scientific world-view, as the logical positivists were, but rather a religious world-view (though on the other hand - just as in the case of Ayer - many aspects of Dummett’s theory, such as the behaviourism implicit in it, do not fit well with a religious world-view either).

Unlike the logical positivists Dummett admitted (in [Dum91, page 11]) that metaphysical realism and anti-realism had content i. e. were meaningful instead of gibberish, and further that the choice between them is not a mere convention, as desperate logical positivists had sometimes suggested (and many of the more extreme relativists do today). Nevertheless, Dummett claimed (e. g. in [Dum91, page 15]) that the content of metaphysical questions, properly understood, is a thesis in the theory of meaning, so
he does in the end attack metaphysics as a separate discipline. Dummett claims (see [Dum91, page 10]) that metaphysically formulated claims are pictures, whose literal content is unclear, and only become clear when they are reformulated as semantical claims. This is not a novel claim; as we have seen, the late Ayer held it, and Gustav Bergmann, though he rebelled against the anti-metaphysical tendency of logical positivism, could not free himself from it. While I argue in this work that the theory of meaning or semantics is essential to metametaphysics, the clarification and justification of metaphysical claims, agreeing so far with Dummett’s approach from the bottom up, I most emphatically do not want to say that the content of metaphysical claims would be semantical. These are two entirely different claims, and could only be conflated by a verificationist. Since Dummett is in some strong sense a verificationist, this may not be surprising; however, since it is semantical research which Dummett claims to provide an answer to the problem of realism, it seem highly premature for him to characterize the problem in a way that presupposes an anti-realist answer to it before he has engaged in the actual research. While I also agree that the content of most traditional metaphysical claims is unclear before a semantical analysis, this of course does not imply that these claims would be metaphorical, or that their clarification by semantical means would turn them into semantical claims any more than the clarification of a claim in the natural sciences by a philosopher of science turns such a natural scientific claim into a semantical claim. While metaphysicians often have to employ pictures and metaphors for pedagogical purposes, so do the researchers in any discipline remote from common experience (including semantics); most metaphysical claims are wholly literal as such, without any need for radical reformulation.

Furthermore, Dummett holds that the theory of meaning is the theory of understanding, so metaphysical claims end up according to him as claims in the theory of understanding, i. e. in the philosophy of mind or even in psychology. This is far from a clearly correct claim. While there clearly are close connections between semantics and the theory of understanding, they can scarcely be simply identified. Identifying them seems like claiming that optics is the same as ophthalmology. It can be argued that it already introduces a commitment to anti-realism concerning meaning. Externalist theories of meaning would typically dispute this claim, as argued e. g. by Panu Raatikainen in [Raa05], as they hold that the meaning of a sentence
consists in its truth-conditions and these are determined by the reference of the terms occurring in it but persons using a sentence need not know what those terms refer to. It seems to me that internalists in some senses of the word need not accept this claim either; even if the meaning of a word supervened on the neural states of its speakers alone instead of on causal chains reaching beyond speakers’ internal states speakers need not be aware of such neural states.

Perhaps the most popular version of the verificationist principle today is Dummett’s manifestation requirement, backed up by his acquisition argument. This requirement says that if a person understands the meaning of a sentence he must be capable of completely manifesting this understanding behaviourally.

---

79 The acquisition argument is based on the idea that a person cannot learn to apply an expression to unobservable entities. The argument is based on a simplistic view of what it is to learn to use a language. As e. g. Noam Chomsky has often argued (e. g. in arguing against Skinner in [Cho59]), the learning of a language cannot be understood as a simple matter of conditioned reflexes. In order to be at all capable of learning a language, a person must be capable of using inductive generalization; from observing a finite number of uses of a language he must be able to infer how an infinite number of sentences are well-formed as well as what meanings the infinite number of sentences all have. I will argue at more length later that induction can lead to knowledge of unobservable entities. If this is true, then since the learner must in any case be capable of using induction, he can use it to make generalizations applying to how the word would apply to unobservable entities, including the mental states of the people teaching him to use the language. Herbert Feigl argued in [Fei58] that if behaviourism is replaced with identity theory, we must return to the view that knowledge of the mental states of others is based on analogy (which is a form of induction).

80 Dummett also appeals to Wittgenstein’s use theory of meaning and his private language argument in arguing for the manifestation requirement. Sometimes these are viewed as additional reasons to accept the manifestation requirement. However, Wittgenstein was (at least at one stage of his development) a verificationist, so the use theory of meaning and private language argument themselves were probably motivated by verificationism, so they cannot count as reasons to accept the manifestation requirement independent of verificationism. Dummett said in [Dum78, page xxxiii] that he had not space to rehearse the argument, and simply recorded his conviction that it is incontrovertible. This is rather high-handed, since (as can be seen in the discussion of the argument [CW12]), there is little agreement even about what the argument is. If Wittgensteinian philosophers cannot even agree on what the form of the argument is, how could they possibly be as sure as they seem to be that it is valid? Since then there have been many arguments against Wittgenstein’s conclusion; e. g. Baker and Hacker and Jussi Haukioja in [Hau04] have argued in favour of solitary rule-following and hence of the possibility of a private language. It seems to me that the rejection of verificationism requires also the rejection of the use theory of meaning, at least if it is understood in so narrow a way as Dummett understands it (though as I will show later, not if it is understood in a wider way as e. g. the later Horwich seems to understand it). Therefore the use theory of meaning in this narrow sense of the word is not consistent with scientific realism. In a wide sense of the word “use” it is indeed plausible that the meaning of expressions supervenes upon
Like the verificationist principle, the manifestation requirement can be understood in stronger or weaker senses. It can be taken to hold either that statements about understanding can be conclusively verified by means of observations concerning behaviour, or just that observations concerning behaviour should be relevant to the determination of the truth or falsehood of statements concerning understanding. Dummett admitted in [Dum78, page xxxviii] that he had had usually concentrated misleadingly on a form of anti-realist theory of meaning where the meaning of a statement is given in terms of what conclusively verifies it, but such conclusive verification is not always to be had. Noting in [Mis95, page 137] this vacillation on Dummett’s part, Misak thought (see [Mis95, pages 139,140]) that Dummett had gone wrong in this admission and that we should stick to conclusive verification, arguing against such philosophers as Anthony Appiah who held in [App86] that the trouble with Dummett’s requirement was that it was stated in terms of conclusive verification. We have seen that there are strong arguments by such philosophers as Berlin, Church, Kaplan, Lewis and Achinstein that the weaker versions of the principle are vacuous, and are so for logical reasons, and if so they must be as vacuous when applied to understanding as when applied to anything else. If this is correct they could not be used to validly support an anti-realist metaphysics any more than opposition to metaphysics. Understood in this weaker sense the manifestation requirement would permit even the understanding of verification-transcendent truth-conditions to be manifestable, though not when understood in the stronger sense.

Dummett argues in my view quite plausibly from the strong reading of the manifestation requirement to a general anti-realism. There have been philosophers - e.g. Mark Quentin Gardiner in [Gar00] and Christopher Peacocke in [Pea86] - who have thought that the argument is in the end their use (though it is not identical with it, as this would destroy the distinction between semantics and pragmatics). However, in this sense of the word, the use of the expressions of a language need not always be accessible to the reflection of a speaker of that language, as Dummett supposes, but can already be evidence-transcendent. The causal relations between expressions, neural processes in their users and the entities the words are used to speak about are partly constitutive of the use of expressions, so a use theory of meaning in the widest sense of the word is not incompatible with a causal theory of reference as is often supposed. However, such causal relations may be evidence-transcendent, and so can the use that they constitute. Just as we cannot even by means of dissection observe the atoms of which our muscles are constituted, so we may be permanently incapable of observing some of the relations, causal and otherwise, between us and the world that constitute the meanings of our words.
not valid, but it seems likely that they implicitly used the weak reading of
the requirement. However, at most Dummett’s arguments show that if we
are to be strong verificationists about statements concerning understanding,
we must be strong verificationists about all statements. Dummett does not
give anyone who is not inclined to be a strong verificationist about anything
any reason to change his opinion; taken as a general argument on behalf
of strong verificationism Dummett’s manifestation argument comes close to
being circular.

There is also the problem that it is far from clear that even a justifica-
tionist theory of meaning could satisfy the manifestation requirement, if it is
taken to demand conclusive verification of claims regarding understanding.
A justificationist claims that we can manifest our knowledge of the meaning
of a sentence by accepting the sentence in some canonical conditions (or as
Dummett says in [Dum06, page 59], when suitably placed), conditions in
which we can verify that statement. However, the problem is how to specify
these canonical conditions, how to tell when we are suitably placed. Most
statements can be justified in the most varying ways in different situations,
and it is not clear that there is anything in common between these ways that
could count as verification conditions that every speaker would understand.

Furthermore, it can be argued that there is no way to specify those
conditions, at least not without lapsing into the kind of semantic holism
Dummett himself rejected. The obvious way in which a statement can be
verified is by it being perceived that it (or something from which its truth
follows) is true. However, contrary to crude behaviourism even a plausible
form of physicalism (such as the identity theory) must agree that percep-
tion requires the occurrence of neural processes in the perceiver. Even if
your eyes are open and in working order (which itself may not be observable
to a typical language-learner, but only to an ophthalmologist), and there
is sufficient light, this does not guarantee that you are seeing anything, if
your nerves are not functioning normally. While such neural processes may
be in principle observable (though even this is not clear), they cannot be
observed in common situations but at most in laboratories with the aid
of rare cerebral scanning equipment (if this counts as direct observation)
or in ethically unacceptable vivisection experiments. Therefore there is in
most situations no way in which people who would ordinarily be said to
understand a statement can be sure that someone perceives that it is true,
i.e. that it can be verified. A child who is learning a language certainly
does not observe the neural processes of those from whom he learns the
language. However, in that case there cannot be any way to conclusively
verify that someone knows how a statement can be verified either. Doubts
about whether we can understand verification-transcendent truth-conditions
are then extensible to doubts about whether we can understand even veri-
fication conditions. Carried through consistently, Dummett’s manifestation
requirement and acquisition argument would result in the absurd view that
we cannot have any understanding at all of the meaning of any sentences
speaking about physical objects.

Like the early logical positivists, Dummett also had a very impoverished
conception of experience, learned from Wittgenstein, excluding introspection
and memory, which led him to troubles with such statements as statements
concerning the past or concerning pain.

Dummett therefore missed all the lessons that he should have learned
from the development of logical positivism, and because of his popularity, a
lot of philosophy has regressed to the point where it was at the the beginning
of the logical positivist movement.

One recent theory used to attack metaphysics is deflationary semantics,
which tries to deflate the metaphysical implications of semantical theories
like the theories of truth and reference. There are curious connections be-
tween deflationism and verificationism, which may be rather surprising as
some deflationists (such as Horwich) have tried to deny them. Already Ayer,
the arch-verificationist, was also one of the first deflationists, defending a re-
dundancy theory of truth in Chapter 5 of [Aye36a]. This might of course
be just a coincidence, but there are more important indications of a tight
connection between deflationism and verificationism: as I will later show at
length in 4.3.1, Hartry Field, one of the most famous deflationists, has mo-
tivated his version of deflationism by the verificationist theory of meaning
in [Fie94].

3.6 Other Miscellaneous Objections to Metaphysics

Ordinary language philosophy has also been used as an objection to meta-
physics (though some philosophers, such as Peter Strawson have tried to
combine ordinary language philosophy with the acceptance of a kind of meta-
physics, descriptive metaphysics, so the incompatibility of ordinary language
philosophy and metaphysics is not obvious).

Ordinary language philosophy was partly inspired by the philosophy of
the late Wittgenstein\textsuperscript{81}, as logical positivism was partly inspired by the
earlier philosophy of Wittgenstein; it was mostly developed in Oxford. Or-
dinary language philosophy concentrated attention instead of syntax (like
the earlier Carnap) or semantics on pragmatics (like the later Carnap). It
stressed the multiplicity of possible uses of expressions and often denied the
possibility of any systematic generalizations, and is in this respect hostile to
metaphysics, whose very definition is based on extreme generality.

Ordinary language philosophers typically argue that metaphysicians use
words outside of the context of their use in ordinary language and that leads
to their sentences being meaningless or in some way out of order. However,
it must be replied that a scientist also uses words in contexts different from
those in which they are used in ordinary language. If the objection of ordi-
nary language philosophers to metaphysics were valid, it would also count
against natural sciences. Therefore no one who thinks that any scientific re-
search at all is legitimate can consider this a valid objection to metaphysics.
This problem in the argument of ordinary language philosophers against
metaphysics is of course similar to the way in which the verification crite-
ron would rule not only the sentences of metaphysics but also the sentences
of sciences meaningless.

Just like logical positivism, ordinary language philosophy is outdated
and should be dead but unfortunately is not. The striving for systematic
generalizations is the very essence of any kind of science - not just science
in the sense of natural science but science in the more general sense of sys-
tematic research or scholarship (what Germans call Wissenschaft). Someone
who denies the possibility of systematicity has no right to pretend to be a
scholar. Ordinary language philosophers often try to justify their view by
pointing out to other uses of language than making statements. However,
not only the making of statements but all important use of language depends
on its relations to the world. We can give a semantic theory not only for
indicative statements but also for commands and questions by considering

\textsuperscript{81}Ordinary language philosophy was also presented (especially by Norman Malcolm)
as developed from the common sense philosophy of G. E. Moore; however, it differed
significantly from Moore’s original philosophy.
their relations to the world. Just as statements have truth-conditions, conditions under which they are true, so commands have obedience-conditions, conditions under which they are obeyed, and a question has conditions with respect to every possible answer to it under which that answer is the correct answer to it. Formulating these conditions involves in all cases assigning interpretations, denotations and intensions and characters and other kinds of semantic values, to expressions occurring in these statements or commands or questions, and these semantic values are entities in the world, entities all of which fall under some ontological categories.

While every systematic reason given for the impossibility of ontology is shown untenable, anti-metaphysical philosophers come up with ever new reasons to declare ontology impossible. Indeed, even a single anti-metaphysical philosopher has often sought several different reasons for opposing metaphysics. As an example, Carnap had at least three different reasons for opposing metaphysics (verificationism, syntacticism and the distinction between external and internal questions in \[\text{Car50}\]) and Ayer likewise at least three (in his case verificationism, common sense and ontological relativism). This strongly suggests that the motivation to deny the possibility of metaphysics is not rational but some purely irrational prejudice that is only rationalised afterwards. We might even, coining a word, speak of an ontophobic attitude.

Some philosophers have attacked metaphysical and ontological problems not as meaningless but as trivial, and have claimed that since the problems are trivial their study is worthless. There are indeed good reasons to suspect that many ontological problems, including many that analytical metaphysicians have discussed at great length, are relatively trivial. However, this does not make their investigation completely worthless. Many of the questions dealt with in many academic disciplines are relatively trivial. Perhaps

---

82I will discuss this pseudo-distinction later in Section 5.9, when explicating Quine’s theory of ontological commitment, since this distinction is, I will argue, based on an erroneous conception of ontological commitment. Carnap did indeed try to show that there was continuity in the reasons of his opposition to metaphysics, for he claimed that the non-cognitive character of what he called external questions in \[\text{Car50}\] was already recognized by the Vienna Circle. However, this is a very problematic claim, for the logical positivists of the Vienna Circle never explicitly formulated any concepts like the concept of an external question that Carnap used in \[\text{Car50}\]. Therefore if they recognized such a distinction this must have been just implicit in their theory; however, since as I will show the distinction is rather obscure, this is also obscure, though I will show that some remnants of verificationism did lie concealed in Carnap’s later distinction.
the extreme example is formal logic, which might be described as in one sense completely trivial, but whose study is yet generally held to be of great importance. Ontological questions are on any construal of ontology no more trivial than logical questions, especially since which some of them indeed are, as I have argued and will later argue further, very closely connected with them.\(^{83}\)

Very few philosophers recommend that we should completely abandon the study of logic even though they hold that logical statements and questions are trivial; while philosophers may differ over why the study of logic is useful or how useful it is, most of them think that it is in some way and at least to some degree useful. For instance, it has been suggested that the knowledge of logical laws is not of any intrinsic importance yet it can used to check the correctness of deductive reasoning used in any disciple whose results are held to be important, and so is of indirect importance. Of course, people can reason deductively even without the development of a systematic theory of logic, but such reasoning is likely to be full of mistakes which can be avoided with the aid of systematic theory of logic. However, others think that even the study of pure logic would be of some limited intrinsic value, even if intuitive deductive reasoning would be reliable enough not to need the aid of deductive logic (which can or course be argued to be a too optimistic view of human cognitive capacities).

If then the study of logic can be useful even though logic is trivial, then if as I argue many ontological problems are trivial either in the same way as logical problems or at least in a very similar way, their study can also be useful in the same way or a similar way as the study of logic. Just as reasoning in any discipline must conform to the laws of logic so the theories of any discipline must be consistent with ontological principles. Ontology might then be of indirect importance for other disciplines. To give just one\(^{83}\)

\(^{83}\)We can distinguish two kinds of concepts of information and informativeness; epistemic and semantic information. Standard theories of semantic information do result in the claim that logically true statements have zero semantic information. A true statement whose semantic information is zero might yet have high epistemic informativeness, if many people do not see that it is true without the aid of artificial methods or even due to some confusion reject its truth. It is often useful to point out such confusions. There are indeed generalizations of the concept of semantical information on which at least some logical statements are semantically informative, such as Hintikka’s surface information; however, some logically true statements lack even surface information. Nevertheless the study of even such statements is not wholly useless, since many of them are surely highly epistemically informative. I think the same holds of many ontological statements.
relatively timely example, computer science has recently made much use of ontology. Computer scientists indeed often use the word "ontology" in a sense completely different from the philosophical sense, referring to any classification as an ontology. However, what many of them call "upper ontology" or "upper level ontology" or "top-level ontology" is really connected to ontology in the philosophical sense as it describes very general concepts that are the same across all knowledge domains, and philosophers such as Barry Smith have contributed to its development. Computer scientists, however, use ontology, even upper ontology, mostly to organize existing knowledge, so the usefulness of ontology for computer science does not require that it would be capable of producing new substantive knowledge of its own.

However, as it can be held that the study of pure logic is of a limited intrinsic value, so it can also be held that the resolution of ontological questions can be of a limited intrinsic value even if it could not be applied to any other subject.

With respect to many metaphysical problems I find plausible a metametaphysical view that David Manley in [Man09, page 4] calls mild deflationism, separating it both from strong deflationism (which is anti-metaphysical) and from mainstream metaphysics. Manley says that mild deflationists admit that there is a genuine dispute at issue, but believe that it can be resolved in a relatively trivial fashion by reflecting on conceptual or semantic facts.

I will argue that many of the metaphysical problems dealt with in analytical metaphysics are relatively trivial. While I cannot of course in this work go through all metaphysical problems and divide them into trivial and substantive ones, I will argue for this in the case of two important and historically prominent metaphysical problems, namely the (basic part of the) problem of

---

84 One advocate of strong deflationism, Eli Hirsch, has also (in [Hir02]) called it a shallow approach to metaphysics; mild deflationism could also be called a mildly shallow approach to metaphysics.

85 Manley gives as an example of mild deflationism in [Man09, page 25] the metaontological theory given by the Neo-Fregeans Bob Hale and Crispin Wright in [HW09]. David J. Chalmers calls this type of view in [Cha09, pages 78,94-99] lightweight realism. I must stress that though I support mild deflationism aka lightweight realism with regard to many ontological problems, this does not imply that I would agree with all of the metaontological views of Hale and Wright; I do agree with some of them but not all. Hale and Wright have been strongly influenced by Dummett, and hold some form of verificationism, which I will argue against.

161
universals and the metaphysics of truth. I want to argue that strong forms of nominalism are just trivially inconsistent with obvious facts while weak forms of realism follow deductively from obvious facts. I will also argue for a weak correspondence theory which takes the truth predicate to be a genuine and complex predicate (and so is opposed to strong deflationism) but does not take truth to be an explanatory concept (as substantivalist theories of truth are by many supposed to do) either and is therefore mildly deflationary.

However, I do not want to hold a mildly deflationary view with respect to all traditional metaphysical problems; I agree that there are some genuinely substantive metaphysical problems, but I tend to be rather pessimistic about the possibility of these problems being solved, at least in the near future. I suggest that the division between those metaphysical problems of which mild deflationism holds and those which are more substantive corresponds roughly to the division between ontological problems and (other) metaphysical problems made by those who distinguish between ontology and metaphysics, starting from Wolff; it corresponds especially closely to the distinction between ontology and metaphysics made by realistic phenomenologists. Of course, even such philosophers have sometimes misclassified questions so that some questions which are not relatively trivial have been held to be part of ontology, so we cannot rely on their authority in the division of problems to ontological and metaphysical.

One attack on metaphysics is directed at metaphysics as an independent discipline. Many people today think that metaphysics is superfluous, since they think that natural science already answers metaphysical questions. Of course, the premises of this kind of attack on metaphysics are entirely incompatible with those of the attack that holds metaphysical questions to be meaningless, since this view holds that they are meaningful questions which physics can answer. If physicalism is true, then physical entities are fundamental entities, so it is suggested that physics already counts as metaphysics in the sense of aetiology, the science of the fundamental constituents of all

---

86 Manley himself mentions such questions as whether there is a God as examples of questions with respect to which people are generally not tempted into deflationism, whether mild or strong. I suspect this also holds of many traditional problems of philosophy such as the problem of the relationship of mind and matter. However, I do not think that it holds of other problems such as the Problem of Universals or most questions relating to the ontology of mathematics.
things. However, this would leave metaphysics as ontology partially independent of physics, as it would examine truths common to both fundamental and derivative entities, drawing both on physics and other sciences.

Also even if physicalism were true, it would yet have to be shown rigorously that it is true, and physics by itself cannot show it. In order to show that for example psychological entities are less fundamental than physical ones (i. e. can be reduced to physical entities, or supervene upon them etc.) one must compare the results and methods of psychology to those of physics. A physicist is not qualified to perform such a comparison simply by reason of his expertise in physics. This task requires someone who views all of knowledge together, synoptically, and this task has traditionally been entrusted to philosophers, specifically metaphysicians.

Besides, metaphysics considers more abstract and general questions regarding fundamental entities than even theoretical physics. Even if it were shown conclusively that all entities or all fundamental entities were physical, there would yet remain a distinction - although in that case a rather subtle one - between examining all (fundamental) entities as physical and viewing them simply as (fundamental) entities. It is no part of the metaphysician’s task to ask such questions as what fundamental particles there are. On the other hand, a physicist as such need not consider whether the fundamental entities are enduring entities or events and processes, or whether they have bare particulars as constituents or are just complexes of properties. A physicist might even consider many such questions too trivial for him to consider.

While metaphysics is intended to be a separate discipline, this does not imply that there would be in practice any sharp line between it and other disciplines such as natural sciences or mathematics, since there are not such sharp lines between any other disciplines; for example, it is pretty impossible to draw a sharp line in practice between physics and chemistry or between theoretical physics and mathematics. The separation of disciplines cannot imply that they would be wholly independent from each other; chemistry is surely dependent on physics, and neither can metaphysics then be wholly independent of the natural sciences. Every discipline has occasionally to encroach on the territories of other disciplines, appealing to results derived in other disciplines and producing results that they can in their turn appeal to. Even if metaphysics is continuous with natural science, as naturalistic
metaphysicians claim, this would not impugn its claim to be a separate discipline. In fact theoretical physicists and metaphysicians occasionally have to consider the same questions, as do metaphysicians and mathematicians, and metaphysicians and general linguists etc. For example, such questions as whether time is linear or cyclic and whether time travel is possible lie in the line between metaphysics and physical cosmology. Even here, however, the approach to these questions may be different; metaphysicians mostly ask whether e. g. time travel is conceptually possible, while physicists ask whether it is actual or nomologically possible.

Of course, the fact that all attempts to prove metaphysics impossible or illegitimate have failed does not show conclusively that metaphysics is possible and legitimate. However, it does offer inductive proof that there are no systematic reasons to think that metaphysics would be impossible which we could discover. In such circumstances, trying to develop metaphysical theories is surely justified even if it is not wholly certain that the attempt has any chances of success.

It may be thought that besides such systematic reasons as the verifiability principle, structuralism, syntacticism and coherentism the opponents of metaphysics had another reason to doubt the possibility of ontology, namely the fact that much traditional ontology had been unsuccessful. It is indeed controversial whether and how much progress has occurred in metaphysics. Personally I think that quite a bit of progress has occurred, but admittedly it is very modest compared to the progress of natural sciences, and it is hard to prove that even modest progress has occurred. Positivists used to point to the lack of agreement after centuries of discussion as a reason for the impossibility of metaphysics. Of course, if this were a valid reason to doubt the possibility of ontology, it would be valid reason to doubt the possibility of all philosophy, since despite the bright hopes of the positivists, later history has shown that no more agreement has been reached in anti-metaphysical philosophy than in metaphysical philosophy. It would also be a reason to

[87] Indeed, agreement is hard to find even in the history of philosophy, though one might think that there we have just to follow ordinary historical methods. However, one could just as easily challenge the possibility of progress in history as in metaphysics. I will later show as an example that there is no more agreement about what Aristotle thought to be the correct answer to the problem of universals than about what is the correct answer to that problem. Obviously philosophers' own opinions about universals affect their interpretation of Aristotle. Philosophers who respect Aristotle typically want to find in Aristotle's doctrine support for their own beliefs and therefore interpret him as having
doubt the possibility of nearly all humanistic disciplines; for instance, the
disagreements between rival schools of psychology, between behaviourists,
psychoanalysts, humanistic psychologists, transpersonal psychologists and
cognitive psychologists, are just as deep as those between rival schools of
philosophy. Therefore metaphysicians are in the same boat with psycholo-
gists and sociologists and literary researchers.

However, mere lack of increasing agreement with respect to any disci-
pline does not do anything to justify the claim that the statements of that
discipline would be meaningless or without truth-value. It does not do it
in the case of ethics any more than in the case of metaphysics. At most it
could provide evidence for the sceptical or agnostic thesis that metaphysical
or ethical questions are too difficult so that it would be impossible for us
to attain justified beliefs in metaphysics. For some reason, today’s philoso-
phers are very unwilling to seriously consider this kind of thesis; one must
be afraid it is just arrogance that blocks this alternative from their sight.
Even so, lack of increasing agreement would not provide any evidence for
the thesis that progress in metaphysics would be impossible in principle,
as Kant claimed would be the case for transcendent metaphysics, metaphysics
dealing with things in themselves. Such a strong thesis would require a sys-
tematic grounding in some ambitious epistemological doctrine like Kant’s,
about whose truth no more agreement is likely to emerge than about the
metaphysical theories whose truth it would pronounce unknowable. Lack of
increasing agreement could at most provide evidence only for the less radical
sceptical thesis that metaphysics would be impossible in practice.

Now I do not think that we should be too sanguine about how much
progress is possible in metaphysics, and this is just because of this kind of ev-
idence. Nevertheless, the scarcity of increasing agreement scarcely provides
sufficient justification for even the claim that all progress in metaphysics
would be impossible in practice, but only for the depressing yet bearable
claim that the probability of progress is rather small, and if any progress
occurs it can only be very slow and modest. I do fear that we can in the
near future gain rigorous knowledge only about the most trivial metaphysical
truths, but even such relatively trivial results are of some worth. The dismal
roughly the same view as they do. On he other hand those who do not respect him often
want to find as big a gap as possible between their views and Aristotle’s. Both motives
lead to the distortion of Aristotle’s own ideas.
prospects of progress scarcely justify us in ceasing to investigate anything but the hard sciences where alone indisputable progress appears; mankind’s desire for knowledge is not satisfied by having knowledge only about physical and chemical questions. If someone is genuinely interested in metaphysical questions, he is justified in investigating it even if the probability of progress is small, so long as it is not completely non-existent.

In fact one important reason for the bad quality of much traditional metaphysics is very clear; metaphysics has typically been used to defend beliefs that had been adopted beforehand, independently of scholarly research, and often for irrational reasons, as at least often in the case of traditional religious beliefs. Defenders of religious orthodoxy have usually not scrupled to make knowingly use of fallacious arguments to defend the beliefs they consider important. On the other hand opponents of traditional religious beliefs have not always if ever been more honest but have also used metaphysics just as a weapon to attack traditional religious beliefs, and have used clearly fallacious arguments for atheism. Very few metaphysicians have been even close to impartial or intellectually honest in their research. Of course the questions of metaphysical would not be so easy even if prejudices could abolished. Many metaphysical and especially ontological problems are so abstract that maintaining a clear idea of what one is talking about is quite difficult (though it would surely not be more difficult than in more rarefied branches of mathematics like set theory or algebra if irrational prejudices did not compound the difficulty). Thus it is no surprise that traditional metaphysical discussion has not led to any kind of agreement.

The situation has of course not changed much in this respect; religious questions and other questions connected with one’s world-view still often arouse vehement passions and philosophers can seldom avoid being swept up in them. Thus I think that an intellectually honest philosophical discussion of such difficult and sensitive questions as the existence of God is nearly practically impossible even today (which does not imply that we should not strive for it). However, when ontological discussion is confined to more abstract questions which are not of much practical relevance, such as the problem of universals, it should be easier to maintain some amount of intellectual honesty, though of course even these questions have some connection to the more substantive ontological questions which have religious and ethical importance, so the maintenance of intellectual honesty is difficult
even in discussing them. Nevertheless, even if the attainment of agreement
is unlikely, this does not make it impossible for individual philosophers to
reach justified conclusions.

I will argue that the legitimacy of ontology and metaphysics is the only
rational conclusion to draw from the failure of the verificationist principle
and other similar principles. While I will not try to decide the question
of the correctness of empiricism as an epistemological doctrine, I will argue
further that what holds of classical verificationism holds also of other initially
plausible forms of empiricism which have been presented as an obstacle to
metaphysics or to specific metaphysical theories. These forms of empiricism
can also be formulated in different ways that are of different strength, and it
then turns out that either they are weak enough to allow for the possibility
of metaphysics or are so strong that they are also block much intuitively
indispensable non-metaphysical knowledge and so are quite implausible. For
example, one such principle which has often today been presented as an
obstacle to the existence of uninstantiated universals has been the demand
that entities of which one has knowledge must affect one causally, often
known as the Eleatic Principle, which I will discuss in 5.3.
Chapter 4

Metaontology and the Theory of Truth

4.1 Correspondence Theory of Truth and Deflationism

In this chapter I will discuss the metaontological implications of the problem of truth, namely the question of what truth is. In current philosophy the most prominent candidates for theories of truth are often thought to be the correspondence theory of truth and the deflationary conception of truth, though epistemic conceptions of truth still have their defenders. A third popular conception of truth is Tarski’s Semantic Conception of truth. There is much confusion over how it is related to the other two theories, correspondence theory and deflationism. There has been much debate about whether it is indeed a version of the correspondence theory of truth as Tarski himself claimed it was or whether such a theory might better be described as a deflationary theory of truth.

This debate over the interpretation of Tarski’s theory of truth is relevant to my dissertation in two different but complementary ways. The first of them is negative; I will use Tarski’s theory to argue against competitors to my metaontological theory, namely against truthmaker theories. I will argue in Section 6.5 that the Problem of Universals concerns primarily the ontological commitments rather than the truthmakers of predicative statements. To prepare the way for that claim I will here argue in Section 4.2.4 that no sufficient reasons have been given for the claim that ontology requires a
truthmaker principle in the first place. One prominent reason, probably the main reason, why many philosophers have been attracted to a truthmaker principle has been that they think that a genuinely realistic metaphysics requires a correspondence theory of truth and a correspondence theory of truth requires truthmaking. I am in favour of a rather strong version of realism; this dissertation argues for realism about universals, and I view such realism as a natural part of a general realistic theory. I have already argued against such anti-realist principles as verificationism. However, I do not think that a truthmaker theory is the right way to defend or develop a realistic theory, even a strongly realistic theory.

It has been questioned whether correspondence theory is necessary or even relevant for realism, as Michael Devitt has done in [Dev84]. However, there are serious problems with Devitt’s view. Devitt begins by saying that no doctrine about truth is constitutive of realism, which is indeed plausible if ”constitutive” is understood in some very strong sense. However, he goes far further than this. He says in [Dev84, §4.2, page 41] that Realism does not strictly entail any doctrine of truth at all. However, he later in [Dev01b, §43, page 46] says quite plausibly that realism requires that truth be absolute, since it is inconsistent with relativistic anti-realism, and if truth were relative, we could use the equivalence thesis (i. e. the T-schema or Convention T) to derive relativism. I cannot see how these two claims could be consistent with each other. This requirement seems to be a strict entailment, quite unlike the abductive links between Realism and theories of truth Devitt dealt with earlier in that section, since the T-schema is a strict entailment. However, absolutism concerning truth is surely a doctrine of truth.

Devitt claims that though this link between realism and truth is close, it is not significant. However, it is in fact very significant, because (all or most) traditional rivals to the correspondence theory imply that truth is not absolute. Neurath’s coherence theory of truth implied that a sentence was only true relative to a theory or to arbitrarily chosen protocol sentences, and the pragmatic theory of truth implies (at least in most of its forms) that truth is relative to the interests of the inquirer. Curiously Devitt does not even mention these two famous theories, but he does mention epistemic theories of truth (with which pragmatic theories do overlap). However, surely epistemic theories of truth strictly entail rather than just abductively
indicate the relativity of truth\(^1\). What is known is always known by someone and therefore knowledge is always relative to a person or a community, and so must according to epistemic theories of truth be truth also. While epistemic

\(^1\)In contradiction to this Devitt claims that even the epistemic theory of truth does not strictly entail anti-realism. He demands an argument from the theory of truth to anti-realism. The problem with Devitt’s characterization of realism is that while Devitt defines realism partly (in [Dev84, §2.2]) in terms of the independence of the world from the mental, he does not provide any explanation of what he means by independence. Nor does he explain what he means by strict entailment. However, one common theory of ontological dependence, the modal conception of independence such as is presented for example by Peter Simons in [Sim87, §8.3, pages 294-301], is that an entity is dependent on another if and only if it cannot exist without the other existing. It follows from this that an entity is independent of another if and only if it can exist without the other existing, i. e. a is independent of b if \(\neg \Box((\exists x)(x = a) \rightarrow (\exists y)(y = b))\). Strict entailment is also often understood in modal terms; a sentence entails another strictly if the entailment is necessary, i. e. \(p\) entails \(q\) strictly iff \(\Box(p \rightarrow q)\).

If we accept these two modal definitions we can give an argument that anti-realism can be inferred from many epistemic theories of truth demonstratively and not just abductively. Devitt says that if \(T_E\) expresses an epistemic conception of truth, then given the equivalence thesis, the epistemic doctrine requires that the appropriate instances of \(s\) is \(T_E\) if and only if \(p\) hold. However, while the epistemic doctrine does require this, it requires more. Surely, the epistemic theory of truth entails strictly that the epistemic conception of truth is strictly equivalent with the pre-theoretic or intuitive notion of truth, as the epistemic conception of truth is supposed to be an analysis or a reduction of the pre-theoretic notion, and even in the latter case it should be just as necessary as Kripke argued the identity Water is \(H_2O\) to be. Therefore it requires that necessarily \(s\) is true iff \(s\) is \(T_E\) if and only if \(p\) hold. However, the equivalence thesis is also commonly taken to be necessary; if \(s\) is taken as an interpreted sentence (i. e. a sentence which is partially individuated on the basis of its meaning) then it is necessary that \(s\) is true iff \(p\) i. e. \(\Box(T(s) \leftrightarrow p)\) (and uninterpreted sentences cannot be true in the first place).

We can infer in any standard modal logic by combining these that necessarily \(s\) is \(T_E\) iff \(p\), i. e. \(\Box(T_E(s) \leftrightarrow p)\), which entails \(\Box(p \rightarrow T_E(s))\). However, this conclusion expresses some kind of dependence of reality on the human mind for many epistemic conceptions of truth, if we accept the modal definitions of dependence and independence. E. g. if \(p\) were the proposition or sentence that the Sun exists and \(T_E(s)\) were Devitt’s own example of an absurd epistemic theory "The Pope affirms that \(s\)”, then we can surely argue that it is impossible for the Pope to affirm anything if he does not exist and have a mind, i. e. if \(o\) symbolizes the mind of the Pope then \(\Box(T_E(s) \rightarrow (\exists x)(x = o))\) and therefore this epistemic theory entails that \(\Box(p \rightarrow (\exists x)(x = o))\), i. e. that the Sun cannot exist if the Pope’s mind does not exist. According to the modal conception of dependence this then entails that the Sun is ontologically dependent on the mind of the Pope and therefore on something mental. Actual epistemic theories will of course instead entail that \(T_E(s)\) entails the existence of humanity as a species or in some cases just of some kind of mind and therefore entail that physical objects like the Sun are dependent on humanity or some mind.

This argument has of course the problem that the modal conception of dependence is controversial (e. g. Kit Fine has argued against it). However, without having any definite conception of dependence we cannot have a definite conception of what realism is, and an indefinite conception of realism is not of much value, so the burden of proof should here fall on those realists who reject the modal conception of dependence to provide a better conception of it.
theories of truth usually explain truth as what can in principle be known, even this must be relative to cognitive faculties of human beings though not to actual evidence. Strongly deflationist theories are not consistent with the absoluteness of truth either; e.g., Quine’s disquotational theory is so close to a coherence theory that it also makes the truth of individual statements and the reference of terms relative to a theory (as clearly seen in [Qui68]).

So a realist is motivated to defend a correspondence theory of truth since he has to preserve the absoluteness of truth. This link between realism and correspondence theory may not be a strict entailment, like the link between realism and absolutism on which it is based, but it is yet quite close and significant. Devitt does not see this because he sees the correspondence theory of truth as a stronger theory than it is; Devitt thinks that a correspondence theory must be an inflationary theory, but I will show that it need not be. However, even if realism left open the choice between correspondence theory and deflationism, yet it would still have huge entailments for the theory of truth if it excluded all the other theories.

Devitt has a maxim, Maxim 3, in [Dev84, page 4] to settle the realism issue before any epistemic or semantic issue. However, he himself admits that it is oversimplified, as realism is a little bit semantic, since according to it the world must be independent of our knowledge and ability to refer to it. Furthermore, Devitt argues in [Dev84, page 70] against phenomenalism (and so on behalf of realism) by appealing to a truthmaker principle, arguing (as such ontologists as Armstrong have also done) that subjunctive conditionals concerning sense-data would need something that makes them true, but phenomenalists cannot provide such truthmakers. In doing this Devitt violates his own Maxim 3, as such a truthmaking principle is surely a semantic principle. Devitt’s appeal to the truthmaker axiom seems fallacious even if the maxim is abandoned for though he argues for an inflationary correspondence theory, he (unlike Armstrong) nowhere argues specifically for a variety of inflationary correspondence theory that would imply a truthmaking principle. Surely the fact that Devitt himself could not follow his own maxim is a reason to doubt how well-advised it is. Also as we have seen

Also problems with subjunctive conditionals are not exclusive to phenomenalists. Even physicalists who subscribe to a truthmaking principle have a problem with subjunctive conditionals, for natural laws are arguably expressed by such.

Devitt labours in [Dev84, §4.4, page 46] the point that realism does not imply any doctrine of truth in the context of the contemporary debate over scientific realism. However, it seems to me that this is the wrong context. The theory of truth is indeed not of
in discussing Ayer’s theory, the opposing party, the phenomenalists, also appeal to semantic claims, as they argue that statements concerning physical objects can be semantically analysed in terms of sense contents, a claim realists like Devitt must disagree with. Therefore semantic claims seem to be in many different ways closely relevant to the debate over realism.

I will argue that a correspondence theory need not be a truthmaker theory, so that even if Devitt is wrong, a realist need not accept a truthmaker theory. Therefore I will argue that Tarski’s theory of truth can provide (if developed further and generalized) an adequate version of the correspondence theory of truth such as a realistic theory requires, but does not have any need for truthmakers, so a correspondence theory of truth has no need for truthmakers.

The second way in which the debate is relevant to the theme of my dissertation is more positive. The theory of ontological commitment is at the basis of my approach to the problem of universals and I will argue (in Section 5.11) that the theory of ontological commitment is highly analogous to the theory of truth. Therefore choosing the right theory of truth is crucial to the development of the theory of ontological commitment, since the theory of ontological commitment should be developed as an analogue of the theory of truth. I think that Tarski’s theory of truth is the best theory of truth so far found and therefore the theory of ontological commitment should be developed analogously to Tarski’s theory of truth rather than some other theory of truth.

It seems that every possible view about the relation between Tarski’s theory and correspondence theory has been taken, both as to its interpreta-

4This implies that Devitt’s main argument against phenomenalism may not be valid. However, this certainly does not imply that we should accept phenomenalism, as there are many other arguments against phenomenalism. Nevertheless, it does weaken the case against phenomenalism; I do not think that phenomenalism can yet be conclusively refuted, though realism concerning physical objects does seem to be a better hypothesis.
tion and as to how to evaluate it. Many philosophers have even at different times taken different views. Many philosophers, starting from Karl Popper, and including famous philosophers like Jan Wolenski (e. g. in [Wol]) and Ilkka Niiniluoto (in [Nii99] and [Nii04]), claim that Tarski’s theory is a correspondence theory (as Tarski himself claimed) and think that Tarski was right in defending a correspondence theory. However, even they might disagree among each other about what kind or version of correspondence theory Tarski’s theory was. E. g. Ilkka Niiniluoto thinks that Tarski’s theory of truth was a strong, classical correspondence theory while Wolenski and Simons think that it was only a weak correspondence theory. Of course such philosophers may yet think that Tarski’s theory is in some way an unsatisfactory correspondence theory. Some philosophers such as Donald Davidson (in [Dav69]) and Hartry Field (in [Fie72]) have initially held this view but then abandoned it; Davidson changed his interpretation of Tarski’s theory as a correspondence theory (see [Dav90, page 302]) without changing his evaluation of Tarski’s theory but Field apparently only changed his evaluation of it. Other agree that Tarski’s theory was indeed a correspondence theory, but think that such a theory is wrong, and that Tarski’s theory must be replaced by something more deflationary or by something more epistemic; many deflationists (such as Horwich and to an extent the later Field) belong to this group and so do such semantical anti-realists as the late Putnam. Still others, like Devitt, Patterson and most truthmaker theorists, think (often for different reasons) that Tarski’s theory was a deflationary theory and therefore wrong, since it is so weak as to be trivial and something stronger is needed, and therefore we need a correspondence theory (or an epistemic theory). There are also those like Soames (in [Soa77, page 18]) and Quine who think that Tarski was a deflationist and was correct in being so, though curiously nearly all of them claim that Tarski’s theory did not represent quite the right kind of deflationism. Finally, there are those who think that Tarski’s theory is compatible with both correspondence theory and deflationism (perhaps including the later Field). There are also more complex and subtle views about the relationship between Tarski’s theory and correspondence theory that would take long to explain properly. Deflationists often seem to vacillate between accusing Tarski’s theory of being too inflationary and attacking correspondence theorists by claiming that they cannot make use of Tarski’s theory because it is not inflationary enough; the
attitude they take towards Tarski’s theory does not always seem to be even formally consistent. Poor Tarski is attacked from the deflationary direction for being too inflationary and from the inflationary front for being too deflationary; he is caught between two extremes like a social democrat caught between communists and capitalists.

The reason for this diversity of opinions is to be found not only in difficulties in interpreting Tarski’s theory, even though there are some such (not very great) difficulties, but more fundamentally in difficulties in agreeing about what correspondence theory and deflationism are (and therefore about how they are related to each other, e. g. what if anything distinguishes them from each other) and in agreeing about what the problem is to which they provide answers. The example of Donald Davidson helps to confirm this view. Davidson apparently did not change his interpretation of Tarski as a correspondence theorist because of any substantive changes in his own theory, but only because he changed his mind about what correspondence theory amounted to. At first he took object-based correspondence theories to be possible but then came to think that a correspondence theory would have to involve commitment to facts, which he consistently detested⁵. Indeed, Tarski was an exceptionally clear writer, so it is strange that there should be many problems about how to interpret his theory; nevertheless, since even Tarski had to use natural languages in presenting the philosophical motivation and interpretation of his theory, he could not be completely unambiguous, and like all humans he made slips, so it is to be expected that small disagreements about how to understand him should arise. However, not only do philosophers disagree about whether Tarski was a correspondence theorist, but there are also disagreements about almost

⁵Davidson seems to have been extraordinarily unclear in formulating the distinctions between various theories of truth. In [Dav86] he defended a coherence theory of truth and knowledge that yet was in his view compatible with a correspondence theory but did not demand that truth should be definable and was therefore also compatible with primitivism. Both coherence theory and correspondence theory on most understandings of them seem to me incompatible with primitivism, though the words can also be used in a weaker sense in which they are compatible with it. Later Davidson came to consider that he had not been well-advised to describe his theory as either a correspondence theory or a coherence theory. However, apparently he never considered that the very phrases ”correspondence theory” and ”coherence theory” might be ambiguous or vague. Clearly this terminological muddle indicates that the very basic concepts of the theory of truth were so muddled in analytical philosophy that extensive work was needed to clarify them before any solution to the problem of truth could be expected - a work that has hardly been completed yet.
any philosopher who has presented a theory of truth that might be called a
correspondence theory or a deflationist theory, whether ancient or modern.
Not only are there disagreements about whether such ancient philosophers
as Aristotle (whose interpretation is notoriously difficult, as we have already
seen) are correspondence theorists or deflationists, but disagreements concern even modern philosophers. E.g. Grover, Camp and Belnap present
themselves as deflationists, but Hartry Field at one time denied that they
were genuine deflationists. These disagreements are connected with dis-
agreements about how correspondence theory and deflationism are to be
defined; there are circulating any number of different definitions many of
which are held to be equivalent by some but not by others.

Like a lot of philosophical disputes, much of the dispute is just based
on mutual misunderstanding and failure to communicate. The correct an-
swer to the question of whether Tarski’s theory of truth is a correspondence
theory or a deflationist theory is, I fear, rather disappointing. Given some
common definitions of correspondence theory it is a correspondence theory
but given other common definitions it is not. The answer about its relation-
ship to deflationism is similar: given some interpretations of deflationism it
is compatible with both inflationism and deflationism, but given others it
is inflationary. Also given some ways of understanding the problems that a
theory of truth is trying to solve Tarski’s theory counts as a nearly complete
theory of truth but given other ways of understanding the problems, it only
scratches their surface or is not even relevant to those problems at all.

Indeed, given suitable definitions Tarski’s theory of truth can be said to
be both a deflationary theory of truth and a correspondence theory of truth;'
more specifically, it can be said to be both a weak correspondence theory and
a mildly deflationary theory\footnote{Even this is slightly inaccurate, for it seems that rather than a binary or trinary
distinction there are rather continua of more and less inflationary theories, and moreover,
theories which are more or less inflationary in different ways. The distinction is really
a matter of degree. However, a division of the non-epistemic theories into three is good
enough for many purposes.}. I will argue that it can be viewed as both (so
long as the Convention T or T-schema is formulated in terms of translation
and is not understood as just disquotationalist) and that so interpreted it
is a very plausible theory of truth, though it requires some generalizations,
which I will try to begin providing. Such a theory is then opposed both
to substantival correspondence theories of truth (such as truthmaker theo-
ries) and to strongly deflationary theories of truth (where most deflationary theories of truth that have been proposed, both pro-sentential theories and minimalist theories, can be seen as strongly deflationary).

Rather than strongly deflating the notion of truth, I will then have to deflate the very debate concerning deflationism. I will argue that the differences between philosophers who call themselves correspondence theorists and those that call themselves deflationists are often less significant than their similarities, such as their opposition to the epistemic theory of truth. Quite often the differences are purely terminological, so the debate comes quite close to being a pseudo-debate. Nevertheless, I think that there are some genuine disagreements (and several of them) at work behind the debate (though they are not so much about truth as about more general semantic concepts such as reference and meaning).

Deflationism is often presented as a new theory of truth, an alternative to traditional theories of truth, both to correspondence theories of truth and epistemic theories of truth such as verificationism or coherentism or pragmatism and to eliminationism about truth. However, there are good reasons to doubt whether it really is a completely novel option in the theory of truth, or whether it is an option in the theory of truth at all. If it were to be a new option this would require that it would be incompatible with all earlier theories of truth, and I will argue that deflationists do not show that it is so, at least on most interpretations of deflationism. I will in fact argue that when we examine deflationist theories carefully it turns out that a deflationist theory, when its possible ambiguities are resolved, either collapses into one of the traditional theories of truth or some kind of combination of traditional theories or turns out to be compatible with several traditional theories, again depending entirely on how it is understood. E.g. Devitt’s deflationary theory threatens to collapse into either eliminativism or traditional correspondence theory.

In order to explain and try to dispel the confusion we can argue that there are several definitions of correspondence theory and several definitions of deflationism which are commonly used and which need not be equivalent though they are commonly taken to be such. There is not one correspondence theory but there are several correspondence theories\(^7\) and there is

\(^7\)It might then be doubted whether it is legitimate to speak at all of the correspondence theory (or of the deflationary theory or of the coherence theory), using the definite article.
not one deflationism but several deflationist theories, and these are not just different ways to implement the same basic idea but different basic ideas which are confused together. Some correspondence theories are consistent with some deflationary theories (or even imply them) but all correspondence theories are clearly not consistent with all deflationary theories.

In order to show this I will in the next sections of my dissertation compare different definitions of correspondence theory and different kinds of correspondence theories and later I will go through different definitions of deflationism and different kinds of deflationary theories.

4.2 Different Definitions of Correspondence Theory and Different Correspondence Theories

4.2.1 What is Correspondence Theory?

The main evidence for Tarski holding his own theory to be a form of correspondence theory can be found in two statements he makes. In [Tar83a, 153] he says the following:

I would only mention that throughout this work I shall be concerned exclusively with grasping the intentions which are contained in the so-called classical conception of truth (‘true - corresponding with reality’) in contrast, for example, with the utilitarian conception (‘true - in a certain respect useful’).

In order to find out whether Tarski succeeded in grasping those intentions we must find out in what sense Tarski used the words "corresponding with reality". It is interesting to note that Tarski here contrasts the classical conception with the utilitarian i.e. pragmatic conception of truth. In his less technical and more philosophical article [Tar44] Tarski repeats the words

However, we can use the expression "the correspondence theory" as shorthand for "the theory containing whatever is common to all correspondence theories" (and similarly for all families of theories of truth). Even when understood in this way, the expression is ambiguous and vague, since the indefinite expression "a correspondence theory" is also ambiguous and vague; nevertheless, it is determinate enough to be useful in initially characterizing the disputes in the theory of truth. More unambiguous expressions are such complex and unwieldy phrases as "the weak object-based correspondence theory", "the strong fact-based correspondence theory" etc., which stand for what is common to all weak object-based correspondence theories, all strong fact-based correspondence theories etc.
"classical conception of truth", adding that this conception is supposed to be the Aristotelian conception of truth. Tarski said there in [Tar44, page 342] that he would like his definition to do justice to the intuitions that adhere to the classical Aristotelian conception of truth. According to him, these intuitions find their well-known expression in the lines of Aristotle’s metaphysics:

To say of what is that it is not, or to say of what is not that it is, is false, while to say of what is that it is, or to say of what is not that it is not, is true.

Directly after this he said that in more modern philosophical terminology, we could perhaps express this by saying that the truth of a sentence consist in its agreement with or correspondence to reality and that for a theory which is to be based upon this formulation the term "correspondence theory" has been suggested.

Tarski then supposed his theory to be a correspondence theory in a sense in which Aristotle’s theory was a correspondence theory but the pragmatic theory of truth is not. We will see later that Patterson thought that the words "correspondence theory" should be used in a sense in which Aristotle’s theory was not a correspondence theory, namely as referring only to theories in which truth plays a serious explanatory role. Naturally Tarski’s theory is not a correspondence theory in this sense, but then he never claimed that it was and it can be doubted if anyone has claimed that it would be a correspondence theory in such a sense.

Early in the 20th century Polish philosophers such as Kotarbiński and Ajdukiewicz generally used such phrases as "classical conception of truth" and "correspondence theory of truth" to refer to something like an Aristotelian theory of truth. Some of them saw the beginning of the classical correspondence theory in the Middle Ages, as in Thomas Aquinas’s definition of truth as the agreement (adequatio) of thought and thing; Aquinas even at least once (in Quaestiones disputatae de veritate question 1 articulus 3) uses the adjectival participle "correspondens” though not the corresponding nominalization ”correspondentia”. For instance, Kazimierz Ajdukiewicz speaks in [Ajd73, page 9] of the classical definition of truth and says that the classical answer to the problem of what is truth is that the truth of a thought consists in its correspondence with reality and then quotes Aquinas’s defini-
tion. Similarly, Tadeusz Kotarbiński in [Kot66, §17, pages 106,107] contrasts the classical interpretation of truth with the utilitarian (i.e. pragmatic) interpretation of truth and says that in the classical interpretation, truly means the same as in accordance with reality. Tarski’s claim that his theory is a version of the correspondence theory and of the classical conception of truth must be viewed against this background.

However, what I think is most commonly called the classical correspondence theory of truth by philosophers today is a version of the correspondence theory of truth that was developed nearly two thousand years after Aristotle’s conception of truth, and which makes essential use of the concept of state of affairs or fact in the account it gives of truth. For example, Michael Glanzberg says in [Gla09] that it was prominent in the early 20th century and its modern history starts with the beginnings of analytic philosophy at the turn of the 20th century. Perhaps most famous examples of such a theory appear in the work of the British Cambridge philosophers Bertrand Russell, G. E. Moore and Ludwig Wittgenstein; however, other early forms of it appear in German philosophers such as the phenomenologists Edmund Husserl and Adolf Reinach in attempts to replace German idealism with a more realistical alternative. Clearly a correspondence theory developed

8Exactly where this correspondence theory originated is a difficult question in the history of philosophy, similar to the question of where the notion of ontology originated. Russell and Moore certainly did not see themselves as creators of the correspondence theory or of the concept of states of affairs; they referred to them as commonplaces of philosophical knowledge. This question is connected with the question of where the notion of state of affairs originated. This is not just a question of who first used the words "Sachverhalt" or "state of affairs" in a technical sense (though some historians seem to understand it in this way and that is also a difficult historical question), since the words are often used to refer to truth-bearers, to true propositions of some kind (especially true propositions with less fine-grained conditions of identity). Rather, the notion in question is that of a non-linguistic complex entity that is essential to the truth of truth-bearers but is not itself a truth-bearer, or at least not just a truth-bearer. Very often the line between truth-bearers and truth-makers is hard to draw and a theory that is presented as an early version of the classical correspondence theory of truth turns out on closer consideration to be a version of the identity theory of truth - a theory that may perhaps in some of its forms be a version of the correspondence theory of truth but can hardly be called a version of the classical correspondence theory of truth. Carl Stumpf has sometimes been suggested as the inventor of the concept of state of affairs, but he was himself asking where the notion came from. It is sometimes said (e.g. by Barry Smith in [Smi89]) and by Artur Rojszczak) that the first occurrence of the conception of a state of affairs (Sachverhalt) in German philosophy was in a work on logic by Julius Bergmann, a follower of Hermann Lotze. Nikolay Milkov has in [Mil02] attributed the origin of the idea (at least in German philosophy) to Lotze himself. Bergmann defined logic traditionally (and of course very dubiously from a more modern perspective) as the art of thinking (Denken) in the sense of judging (Urtheilen), which he says is also the art of knowing or knowledge (Erkennen).
two thousand years ago and a correspondence theory developed a century

Bergmann said in [Ber87, page 2] that we understand by knowledge such thinking that what is thought in it corresponds (ubereinstimmt) with a state of affairs (Sachverhalt), i. e. which is true (Wahr). This is of course an extremely dubious definition of knowledge (for reasons Bergmann himself acknowledges later and tries to explain away, most importantly that it leaves out the element of justification in knowledge). However, it may well be a very interesting anticipation of what is commonly called the classical correspondence theory of truth. Bergmann says so little about states of affairs and truth, however, that it is yet not quite certain that he is not defending an identity theory of truth rather than a fact-based correspondence theory. The case for attributing the notion of state of affairs to Lotze seems weaker to me; Lotze does use the term "Sachverhalt" a few times, but does so rather casually, so it is not clear that the word is at all a technical term for him. Milkov admits in [Mil02, §6, page 446] that Lotze only uses the word twice in the work on logic in which he according to Milkov introduces it; in the first of these passages it is clearly used in a non-technical sense, and only in the second can we see any connection with deep philosophical questions, especially questions concerning truth. Lotze's works, unlike Bergmann's have been translated into English, but the word "state of affairs" does not occur in the translations of these passages. "Erkenntnis des jedesmaligen Sachverhalts" in [Lot12, §138, page 169] has been translated in [Lot88, §138, page 181] as "special knowledge", which is rather loose but fits the informal context; in the latter passage [Lot12, §327, page 534] the word has been translated more exactly in [Lot88, §327, page 235] as "state of things", which could indeed serve as a good synonym for "state of affairs". Let me quote this passage:

Nothing is therefore left of this inexact comparison except the conviction that even the mere perception of a given state of things as it really is, is only possible on the assumption that the perceiving subject is at once enabled and compelled by its own nature to combine the excitations which reach it from objects into those forms which it is to perceive in the objects.

The context is a discussion that we would today call epistemological (regarding the a priori preconditions of knowledge), though Lotze viewed it as a part of logic (in the very wide sense of the word German philosophers of his day used). The inexact comparison Lotze refers to is the comparison of knowledge to a copy of this object; since correspondence theory is often called a copy theory (although inexacty, as Lotze at least recognizes), Lotze seems to be arguing against a correspondence theory here, not developing one. There is nevertheless here a notion of correspondence whose second member is a state of affairs, though its first member is a perception, not a common truthbearer such as a sentence or a judgment or a proposition. Perceptions would not ordinarily be called true or false; however, the relations of correspondence between perceptions and their correlates and between judgments and their correlations seem very similar, so if one should support a correspondence theory of truth, one could well call perceptions that correspond with their objects or correlates also true in an expanded technical sense, so one can see an implicit correspondence theory of correct perception in Lotze. In any case, Lotze does not use the term "Sachverhalt" in an explicit (even though rather casual) definition of truth as Bergmann does. On the other hand Lotze uses besides "Sachverhalt" also the word "Thatsache", which could be translated as "fact" and understood as synonymous with "state of affairs", which Milkov does not note, though it would give as good a reason to hold Lotze as anticipating the modern theory of states of affairs as his use of the word "Sachverhalt". In any case, however, consistently with Lotze's attacks on a copy theory of knowledge Milkov himself attributes to Lotze in [Mil02, §2, page 439] an identity theory of truth rather than a classical correspondence theory, so even if Milkov's interpretation were correct and Lotze had a concept of a state of affairs, Lotze's theory is probably not a classical fact-based correspondence theory of truth.

180
or two ago are likely to differ widely.

Tarski was of course to some extent aware of and influenced by the modern classical correspondence theory of truth, and Tarski does also give a formulation which even makes use of the notion of states of affairs. However, Tarski says that all of these formulations can lead to various misunderstandings, since none of them is sufficiently precise and clear, and adds a very interesting qualification; according to him this applies less to the original Aristotelian formulation than to the others.

This raises an important question; why does Tarski say that the original Aristotelian formulation is more precise and clearer than the more modern ones? He must surely have had a reason for saying so. Marian David says in [Dav09] that Aristotle’s definition offers a muted, relatively minimal version of a correspondence theory. She adds that for this reason it has also been interpreted as a precursor of deflationary theories of truth. I suggest that one likely reason why Tarski said that this definition was more precise and clear was just because it is so minimal. Tarski probably wanted to make clear his commitment to the most basic ideas of a correspondence theory, but wanted to avoid commitment to any stronger version of the theory. Tarski might have thought that reifying states of affairs as genuinely existing entities was one of those misunderstandings to which the formulations could lead. If this was Tarski’s intention then he acted wisely, since we will see that his definition does not support a stronger theory, such as Moore and Russell (and Husserl and Reinach and many other philosophers) developed at early 20th century. Especially Aristotle’s definition does not make any use of the notions of truth-making or facts, as I will show in Section 4.2.3 of this work, and this suggests that Tarski wanted to avoid making use of them.

While Tarski’s theory is not ontologically neutral but implies an ontology of sets, it does not imply an ontology of facts to which sentences (or beliefs or assumptions or propositions) would correspond in a literal sense, since Tarski does not quantify over such entities when actually developing his theory. Tarski mentions facts only in a brief preliminary remark he himself says is not sufficiently clear, so quantification over facts is not part of Tarski’s theory proper.

This does not imply that there would be no need to accept the existence of facts even if Tarski’s theory is sufficient as a theory of truth. Even if
facts are not needed in the theory of truth, they may be needed for other purposes, for instance as the members of the causal relation. There can be arguments for facts based on the notion of ontological commitment instead of the notion of truthmaker. For instance, some sentential anaphora in bits of natural language discourse we hold to be true seem to involve quantification over facts. I will sketch the argument that could be given for the existence of facts from ontological commitment. As has often been pointed out, sentential anaphora are similar to variables for which sentences can be substituted, so they seem to involve quantification over something. If the quantification involved in such sentential anaphora is objectual, then it must be quantification over propositions or facts, and in some cases (e. g. when saying "The earth shook. It caused the statue to fall."), since propositions are not taken to be causes of spatio-temporal events like a statue falling) it seems that it cannot be quantification over propositions so we have reason to think that it is quantification over facts.

4.2.2 Fact-based and Object-based Correspondence Theories

Wolfgang Künne distinguishes in [K03, §1, page 5] between fact-based and object-based correspondence theories of truth, and Marian David follows him in [Dav09]. Künne argued in [K03, §3,1, pages 93-112] that the earliest correspondence theories such as Aristotelian theories of truth and Kant’s theory of truth were object-based correspondence theories. Such earliest correspondence theories did not suppose that the entities truth-bearers corresponded with were facts. As to their positive nature, this was more obscure, but Künne concluded in [K03, page 108] that according to such theories a mental or verbal predication was true if the predicate fit the object it was predicated of. In the case of a verbal predication the object of the

---

9I will later in Section 4.2.3 examine in more detail whether this interpretation of Aristotle’s theory of truth is correct.

10How is the word “object” to be understood in the characterization of object-based correspondence theories? It must not be understood as referring to any specific ontological or semantical category, as in Frege’s distinction between objects and functions or in the distinction between objects and events. Even if Frege’s controversial distinction were correct, functions could still be objects in the sense in which object-based correspondence theories make use of the word. Künne may seem to contradict this, when he says that paradigmatic elements of the right field of the correspondence relation thus understood are material objects such as mountains and people and when he distinguishes event-based correspondence theories from object-based correspondence theories. However, telling what
predication was typically the grammatical subject of the sentence. According to David this actually involves two relations to an object: (i) a reference relation, holding between the subject term of the predicative judgment and the object the judgment is about (its object); and (ii) a correspondence relation, holding between that object and the predicate of the judgment. However, there are really two more relations at work, (iii) a reference (or denotation) relation between the linguistic predicate and the property or set it refers to and (iv) a correspondence relation between the linguistic subject of the judgement and the property that its referent is predicated of.

Of course, we could equivalently say that a predication is true if an object fit the predicate it was predicated of. We could then equally well take the correspondence to consist of a relation between the subject term and the denotation of the object term. The great limitation of this theory, of course,

are the paradigmatic examples of some class does not tell what the limits of this class are, so he need not mean that all objects would have to be material enduring objects like his examples. In fact Künne admits in [Kö3, page 100] that such things as Helen’s beauty (a property) and the fading of her beauty (an event or process) count as objects in Aristotle’s theory, and therefore presumably in object-based correspondence theories generally. However, he still says that anything denoted by a singular term is an object; this seems too restrictive since I see no reason why entities denoted by predicate terms or expressions of other syntactic categories could not be objects as well (especially given that the same entity can be denoted by expressions from different syntactic categories, as I will argue later to be the case). Whatever he thought, it is in any case clear that the objects in question are in the original version of the theory just the objects of the ideas or presentations which a belief or judgement or assumption which is taken to be a truthbearer contains, whether as subjects or predicates. Therefore in the case of a revised sentential version of the object-based correspondence theory such as I argue Tarski’s theory to be they have to be the objects of the terms (whether singular or predicate terms) a sentence or statement contains, i.e. the entities they designate or have as extensions etc. It seems that entities of any ontological category could be objects in this sense, be the subjects of predications or predicated of such subjects or designated. Events and even facts can be taken to be objects in this wide sense if they happen to exist, at least in the case when a sentence speaks explicitly of events and facts (as for example this very sentence does); what differentiates object-based from fact-based and event-based correspondence theories is that according to the former we do not need facts or events to explain the truth of every truth-bearer (or even every contingent or atomic truth-bearer). However, not all objects in every sense of the word should be allowed among the objects in question, for instance objects in the Meinongian sense, such as the round square, should not be allowed among the objects in question; therefore it might have been better to speak of entity-based correspondence theories than of object-based correspondence theories.

Using such modern notation of as I use elsewhere in this work we can say that if \( \|\alpha\|_g \) signifies the denotation of an entity \( \alpha \), specifically a linguistic predicate) relative to assignment \( g \) where \( \alpha \) is an individual variable (of the metalanguage) denoting in this case a predicate variable of the object language, then the correspondence relation would be the dyadic relation \( (\lambda\alpha)(\lambda\beta)(\alpha \in \|\beta\|_g) \).

11 We could also (to anticipate suggestions I will develop at length later) take the correspondence relation itself to be the four-place relation \( (\lambda\alpha)(\lambda\beta)(\lambda\gamma)(\lambda\delta)(\alpha = \|\beta\| \land \|\gamma\| = \|\delta\|) \).
is that it presupposed that all sentences were of the subject-predicate form.

It seems to me that Tarski’s theory can be viewed as a modernized version of object-based correspondence theories, even though it cannot be viewed as a fact-based correspondence theory or a modernization of such. Davidson saw this in [Dav69], even though he did not use the phrase ”object-based correspondence theory”, yet his basic notion of a correspondence theory without facts was essentially similar. However, later he seems to have thought for some reason that the very notion of a correspondence theory without facts did not make sense. Tarski himself may have been aware (even if dimly) of this and this may have been why he said that the Aristotelian formulation is more precise and clearer than the modern ones. Of course, Tarski is not very precise himself in saying that the difference is that the modern formulations would be less clear; that is clearly not the only difference, since at least the second modern formulation has on any natural interpretation stronger implications than the Aristotelian formulation Tarski quotes (though as I show elsewhere there are other statements made by Aristotle that may have equally strong implications).

Of course, there are significant differences between Tarski’s theory and traditional versions of object-based correspondence theories; Tarski does not think that truth-bearing items (which with him are sentences) would always have subject-predicate structure. Only atomic sentences have a structure like that, and even in their case it is a predicate-arguments structure rather than a subject-predicate structure. However, even in the case of complex sentences it is a relation to objects denoted (whether absolutely or with respect to a sequence or assignment) by the terms of those sentences that determines the truth of the sentence.

Künne says in [Kö3, page 5] that what is essential to all correspondence theories of truth, both object-based and fact-based, is that truth is a relational property and the implied relation is not one in which truth-bearers

\[ \delta \land \alpha \in \gamma \].

In fact this assumption, that truth is a relational property is problematic for me since I want to find a metaontological criterion for ontological claims that would help me to decide the problem of universals, and such a criterion should have to be independent of any substantive solution to the problem of universals. Therefore I cannot assume that truth is a relational property as I cannot assume that it is a property at all since I cannot assume that properties exist, for this might be begging the question against nominalism. Besides, as I will show later, even if we find out in the end that there are some properties, there may be special reasons to think that truth cannot be a property (for instance, because there may be reasons to think that there are no relational properties or because this might lead
stand to other truth-bearers.

This definition of a correspondence theory was anticipated by Bertrand Russell in the third of the requisites he gave in 1912 to a satisfactory theory of truth; this requisite was very similar to Künne’s definition of a correspondence theory, though it was not exactly the same and is not quite equivalent. Russell said in [Rus12b, page 189]:

But, as against what we have just said, it is to be observed that the truth or falsehood of a belief always depends upon something which lies outside the belief itself . . . Hence, although truth and falsehood are properties of beliefs, they are properties dependent upon the relations of the beliefs to other things, not upon any internal quality of the beliefs.

Of course, Russell’s definition can be modified as Künne does so that the idea that the truthbearers are beliefs (which is an earlier one of Russell’s requisites and is presupposed by Russell here) is removed from it and it becomes neutral with respect to the question of what truth-bearers are, though even so it is not quite equivalent with Künne’s definition14. Also it

to the Liar Paradox). However, Künne’s point can be reformulated so that we do not have to assume the existence of properties; e. g. we can say following Devitt in [Dev01b, page 580] that for correspondence theorists the truth term is a one-place relational predicate and add for the sake of clarity that it must be this not only in surface syntax but also in logical form. Devitt himself did not think that this was enough for a correspondence theory of truth, since he thought that a correspondence theory of truth would have to hold truth to have an explanatory role; however, I will argue later that for historical reasons this cannot be part of the definition of a correspondence theory. Devitt also added that correspondence theorists think that the truth-predicate has the standard sort of semantics of such predicates; however, this assumption is problematic, since it is not clear that there is any standard semantics for one-place relational predicates, but it may be that different one-place relational predicates generally have very different semantics.

14Clearly it makes as much sense to speak of entities being external to sentences - at least in the case of token sentences or utterances - as to speak of them being external to beliefs (where again externality must mean externality to token beliefs, belief acts or states of believers, not externality to belief types, whose meaning is less clear). However, in the case of substantially conceived propositions, which are generally supposed to be outside time and space, it is not so clear what externality might mean. We can still make sense of it most plausibly in mereological terms by taking externality to mean the lack of common constituents or more weakly the existence of a non-common proper constituent. However, at least a Russellian proposition cannot in this sense be external to the entities on which its truth depends, since if that entity is a fact then it shares its proper constituents with the proposition and if those entities are the entities which the proposition is about then those entities are the (immediate and therefore proper) constituents of the proposition. This, of course, is as it should be; while Russell did not make clear what he meant by externality, he gave this definition after he had abandoned the idea that Russellian propositions were

185
is surely an exaggeration to say that the truth of the belief always depends on something external; you can have beliefs about beliefs, and if Othello believed that he is believing something the truth of this (very rare) kind of belief would not depend on anything external. In the sentential case it is surely possible to have sentences like "This sentence is longer than two words."; in this case the truth of the sentence depends on the sentence itself. It is better to say more weakly that the implied relation is sometimes or often a relation to something other than a truthbearer, whether a belief or a sentence.

This definition (so modified) appears to go beyond the famous Aristotelian formulation I quoted before, though it is consistent with it, and leads us towards a slightly (very slightly) more substantive theory of truth. If the Aristotelian formulation is considered to be consistent with some declarative of truth and had even begun to doubt their existence. In fact even if we do not demand that the entities on which the truth of a truthbearer depends have to be external to it we still cannot reconcile a theory of propositions which takes them to be Russellian with the view that propositions are truth-bearers and with a fact-based correspondence theory, if we assume mereological extensionality (as is standardly done in mereology), since in this case as the Russellian proposition and the fact share their proper constituents, they must according to mereological extensionality be identical. Now it is not clear whether Russell himself was motivated by this kind of considerations, since he rather appeals to the unintuitiveness of objective falsehoods (which is a far weaker argument since other philosophers have not found objective falsehoods unintuitive at all) and Russell of course did not have any explicit theory of parts and wholes i.e. mereology and since Russellian logic was to a large extent intensional he might not have accepted mereological extensionality even if the option had been presented to him. However, many later philosophers have surely noticed at least dimly this inconsistency and been motivated by it in their theories of truth to reject either facts or Russellian propositions. The common argument that true Russellian propositions cannot be distinguished from facts that would make them true appeals tacitly to mereological extensionalism; the only reason to think that facts and true Russellian propositions would be identical is that they have the same constituents and according to mereological extensionalism entities with the same constituents (or rather an identical partition) will have to be the same. However, many important ontologists have rejected mereological extensionality, for instance because such rejection solves problems associated with material constitution even though it is at the price of a bigger ontology. If we reject mereological extensionality and adopt an intensional mereology, where two different entities can share a partition, then we can even reconcile a Russellian theory of propositions and the view of propositions as truth-bearers with a fact-based correspondence theory; we can then take a true Russellian proposition and the fact that makes it true to have the same constituents (i.e. share a partition) but be different since they are put differently together to form different wholes. Whether such a combination is plausible is a different question (though there are many plausible arguments against mereological extensionality), but at least it is possible, and of course if your conception of propositions is not Russellian (but e.g. Fregean) then you can have a correspondence theory for propositional truth with even less cost. So the options in the theory of truth are very wide indeed (and since truth is not the main topic of this work I must leave them so wide open).
flationist theories which are held to be inconsistent with a correspondence
tory (both by themselves and their opponents), then this Russellian def-
imination might formulate the minimal genuine correspondence theory. It is
indeed rather weak, so some might call it also deflationary; however, it is
by no means completely trivial since it has been denied by most deflation-
ists. Many of the more extreme deflationists, such as most prosententialists
(like Dorothy Grover and Nuel Belnap) have denied not only that truth is
a property-ascribing predicate, but even that it would be really a predicate
at all. Even more moderate deflationists such as Paul Horwich, though they
accept that the truth term is a predicate and even that truth is a property,
deny that truth is a complex relational property or even that the truth term
would be a relational predicate\textsuperscript{15}. Therefore this is already a somewhat
controversial principle and substantial at least to a minimal degree.

It is important to notice that this definition leaves a lot of leeway for
what kind of relation the implied relation may be\textsuperscript{16} - and what kind of
entities have to be in its range - so it leaves room for a very great variety
of different correspondence theories of truth. Russell inexplicably went on
to say that this requisite, the third in his theory of truth, leads us to adopt
the view - which he says has on the whole been the commonest among
philosophers - that truth consists in some form of correspondence between
belief and fact. However, if Russell means by "leads" anything like "implies
logically" or "entails" and uses the word "fact" in any technical sense he
commits a clear non sequitur here. This requisite as he initially formulated
it by no means implies that truth consists in correspondence between belief
and fact, only that it consists in some relation between belief and something
external to it. Russell’s claim is at most true if "leads" means something
like "suggests" but in this case the suggestion can be resisted and probably
should be resisted given Russell’s own commitment to Occam’s razor\textsuperscript{17}. It

\textsuperscript{15}On the other hand, Russell’s formulation is (like Künne’s) clearly inconsistent with
the inflationist view of truth that would take it to be sparse property, since it implies
that truth is a relational property, and sparse properties are by definition intrinsic, i. e.
such that they (or more rigorously expressed their instantiation) cannot depend on any
relations of their possessor to other things.

\textsuperscript{16}It is not or at least need not be dependence - Russell spoke of dependence, but so that
according to him the relation is something on which truth depends, the relation itself is
not dependence.

\textsuperscript{17}Of course, you could say that Russell does not in this passage use the word "fact" in
any technical sense here. However, Russell later in any case makes use of a very technical
notion of fact which is associated with a substantive theory, which says (see [Rus12b, page
is very significant that though Russell’s theory of truth is a paradigmatic
fact-based correspondence theory, yet Russell’s initial characterization of the
motivation for the correspondence theory of truth does not mention facts
or logically imply that correspondence would have to be correspondence to
facts. This shows that object-based correspondence theories are a viable
option and do not just constitute changing the subject.

In fact, this definition does not even imply that the implied relation could
be naturally described as a correspondence relation, even though Künne is
not willing to go this far; he says that you cannot have a correspondence
theory without correspondence, but his own definition of the theory implies
otherwise.

Whether either elimination, whether of propositions or of facts, is tenable
in the end is of course highly doubtful, as there are all kinds of difficulties in
implementing either proposal and all kinds of relevant considerations (e. g.
in the case of facts the question whether the theory of causation needs facts
and the semantics of sentential nominalizations and sentential anaphora and
the phenomenology of perception). In fact while I think it unlikely that
either elimination can in the end be carried out, I think that the prospects
for eliminating propositions are in the end, contrary to what Russell thought,
worser than those for eliminating facts. However, when it comes to the
basic principles of the general theory of truth, we can initially do without
either. Wolenski considers this kind of Russellian correspondence theory
with a more than two-place relation briefly, but he rejects it because it has
non-standard truth-bearers; however, it is possible to modify this kind of
theory so that it has different truth-bearers that Wolenski might consider
more standard (even though I rather doubt whether there really are any
standard truth-bearers, so great is the diversity of opinion with regard to
truth-bearers), and I will argue that in fact this is what Tarski implicitly
did.

Künne himself ultimately rejects all correspondence theories of truth in

199) that:

Whenever a relation holds between two or more terms, it unites the terms
into a complex whole.

Facts are such complex wholes, and Russell’s statement taken at face value implies that
there are lots of them. Therefore it is likely that Russell is already presupposing this
technical notion and this theory, and even if he is not this leaves it mysterious why (and
with which justification) Russell later introduces that technical notion.

188
favour of a theory he calls the Modest Account of Truth which can be fairly
called deflationist (even though he does not like that word). However, it
seems to me that this is only because though he resurrected the notion
of object-based correspondence theories of truth, he understood them too
narrowly. He did not see that object-based correspondence theories can be
modified considerably from the admittedly primitive Aristotelian version. I
will in the last section of this chapter show how to develop a new version of
the object-based correspondence theory on the basis of Tarski’s theory that
escapes problems in Tarski’s theory.

4.2.3 Aristotle’s Theory of Truth and Aristotelian Theories
of Truth

While Aristotle’s famous definition suggests a weak correspondence theory
of truth, yet this definition is not the only passage where Aristotle speaks
about truth. Aristotle’s full theory of truth clearly goes beyond this def-
inition. However, since Aristotle’s writings are rather obscure to modern
readers it is extremely controversial what kind of theory it might have been
and in what way it goes beyond the minimal theory. Nevertheless, clarifying
this is essential to understanding the nature of correspondence theory since
as I have argued, Aristotle has often (e. g. by classical Polish logicians)
been taken to be a paradigmatic correspondence theorist, so that a corre-
spondence theory in one common sense of the word is by definition the kind
of theory of which Aristotle’s theory is an instance.

Paolo Crivelli has presented probably the most thorough discussion of
Aristotle’s theory of truth by any modern commentator in [Cri04, page 130];
this is clearly relevant to the subject we are discussing. One possibly du-
bious feature about Crivelli’s interpretation is that Crivelli gives a strongly
realistic interpretation to Aristotle’s theory; he thinks that Aristotle is a
realist about universals as well as states of affairs. While this interpreta-
tion is plausible, less realistic interpretations of Aristotle’s theory are also
equally plausible as I will show later in this work (e. g. Aristotle could also
be viewed as precursor of trope theory)\footnote{It is quite dramatic to contrast
Crivelli’s interpretation with a strongly anti-realistic interpretation such as
that of Gregory Salmieri in [Sal10]. Seeing how much these two versions of
Aristotle differ, one is made highly sceptical about the possibility of our ever
attaining any reliable interpretation of Aristotle’s theory.}. Not being an expert on ancient
philosophy, I must rely to a great extent on Crivelli’s interpretation, which I will try to relate to the modern discussion.

The question whether Aristotle had a theory of facts as truthmakers is extremely controversial. David says that Aristotle sounds much more like a genuine correspondence theorist in the Categories (12b11, 14b14), where he talks of "underlying things" that make statements true and implies that these "things" (pragmata) are logically structured situations or facts (viz., his sitting, his not sitting). The question of how to interpret or even translate such passages in Aristotle’s texts is a very difficult one. Aristotle has a word, "πράγμα", "pragma", that is occasionally translated as "state of affairs". However, the word is very obscure and translating it as "state of affairs" already involves a lot of controversial interpretation. It has also been translated more literally as just "thing" or "object", as we have already seen that Künne translated it. Aristotle’s theory of pragmata might be easily mistaken for a precursor of a fact-based correspondence theory with states of affairs as truthmakers such as Armstrong’s theory. For example, Barry Smith seems to make this kind of assumption in [Smi89], as he says that traces of the Sachverhalt concept are discoverable by hindsight already in Aristotle in those passages where Aristotle speaks of the pragma as that on which the truth of the logos depends, and distinguishes states of affairs from propositions. However, while the exact interpretation of Aristotle’s theory is a very complex question, it can relatively safely be said that it is not a straightforward anticipation of modern fact-based truthmaker theories.

Such an interpretation faces two serious problems. First of all, it is not clear that Aristotle is wholly realist about these states of affairs. Secondly and even more importantly, Aristotle says in Metaphysics (1024b18) that these things or states of affairs are themselves true or false. Because of

---

19 The word in Aristotle’s Greek which most closely corresponds to the English words "beings" and "entities", which express the object of ontology, is "οντα", "ontas". It is not wholly clear how the two concepts expressed by the two words, "ontas" and "pragma", are related; they seem sometimes to be synonymous, like the (probably roughly) corresponding Latin words "entes" and "res", but there do seem sometimes to be differences of meaning. Gregory Salmieri argues in [Sal10, page 19] that "pragma" is not an ontologically loaded term for Aristotle at all, but refers to the object or subject matter of a thought (or other mental state). It would then be like Brentano’s and Meinong’s German term "Gegenstand". If Salmieri is right then the meaning of the words would differ dramatically, and this would of course count against Aristotle’s theory of truth being a truthmaker theory as well as against a realistic interpretation of Aristotle’s theory of universals. I am not at all certain if he is right, however.
this these states of affairs are more plausibly taken to be truth-bearers such as propositions rather truth-makers\textsuperscript{20}. Indeed, it can even be argued that Aristotle thought that these states of affairs are primary truth-bearers, as Crivelli does in [Cri04, pages 56-57].

Crivelli says in [Cri04, page 130] that Aristotle’s theory does not qualify as a correspondence theory on the conception according to which an assertion is true just in case there is some fact to which it corresponds, since Aristotle’s theory does not mention facts. Crivelli says that he uses the word “fact” in the Russellian sense (apparently referring to the phase or Russell’s philosophical development shown in [Rus12b]); every fact, by its very nature, is true (i.e., obtains). This seems to be close to the sense in which such truth-maker theorists as Armstrong use the word, so we may interpret Crivelli as implicitly saying that you cannot find a theory of facts as truth-makers in Aristotle (even though truth-maker theorists would be unlikely to speak of facts being true, though they would speak of them obtaining).

Perhaps surprisingly, Crivelli yet thinks that Aristotle had an ontology of states of affairs. The phrase ”state of affairs” is of course very often (e.g., by Armstrong) used as a synonym of ”fact”. However, Crivelli uses it in a different sense, which seems to be closer to the way many philosophers use the word ”proposition”. In fact, these states of affairs which are not facts or truthmakers seem to be just like Russellian propositions, since according to Crivelli’s interpretation they are composed of individuals and universals.

Of course, many other modern philosophers have used the word in this sense; for example, Roderick Chisholm, Alvin Plantinga and John Pollock. Chisholm in [Chi70] took propositions and events (understood as recurring)

\textsuperscript{20}On the other hand, Aristotle also says in Metaphysics (1027b26-28) that falsehood and truth are not in objects (pragmata) but in thought (διάνοια, dianoia), which seems to contradict his claim that objects (pragmata) are true of false. In fact, the context of the latter claim also provides a possible key to resolving the contradiction, since Aristotle there distinguishes between several senses of being, and apparently the things that are true or false, whether they are propositions or states of affairs, are not (do not exist) in the strict sense. It is not wholly clear what this means, but one way of understanding this might be that these states of affairs or propositions are just logical constructions out of thoughts (entia rationis as medieval scholastics or such Austrian philosophers as Brentano said). We can suggest that this less than strict sense of existence corresponds to what would in today’s philosophy be called non-objectual quantification. This would leave us with an anti-realistic account of states of affairs or of propositions. Of course, Aristotle can also be interpreted so that all of his senses of existence correspond to different kinds of objectual quantification, which would give us a realistic theory of states of affairs or propositions.
to be a subspecies of states of affairs; the difference Chisholm makes between propositions and states of affairs might also be described as a difference between tenseless (eternal) and tensed propositions. Furthermore, the states of affairs into which moderate modal realists such as Alvin Plantinga in [Pla74, page 44] want to reduce possible worlds are also similar to Russellian propositions. Plantinga identifies possible worlds with maximal states of affairs. While Plantinga leaves in [Pla74, page 45] open the question whether states of affairs and propositions are identical (referring to Chisholm’s view), he does say that they correspond to each other21. Plantinga also says that every world (maximal state of affairs) exists in every world but does not obtain in more than one. Plantinga says in [Pla74, page 47] that obtaining or actuality for states of affairs is like truth for propositions. If the actual world $\alpha$ would not have obtained, there would till would been such a thing as $\alpha$. This clearly differs from how Armstrong and his followers think of states of affairs and obtaining; according to them, if a state of affairs (including the totality state of affairs which in a way corresponds to the actual world) had not obtained, it would not have existed, i. e. there would not have been such a thing as that state of affairs22.

One modern philosopher specializing in the theory of truth who actually explicitly states that propositions and states of affairs can be identified is Marian David in [Dav94, page 46]. David adds that the obtaining of states of affairs can then be identified with the truth of propositions. However, David qualifies this in [Dav94, page 47] by saying that when the representation theorist spells out her theory, she may find that there are additional roles that need to be filled, which may motivate some redistributions of her terminology. The representation relation may according to David have to be explained (I would say analysed) in terms of a further notion of content for which the term ”proposition” seems appropriate. Different sentences may express different propositions that determine the same state of affairs, so

---

21Plantinga distinguishes from each other a world as a maximal state of affairs and the book on W which is the set of propositions that are true in that world; however, it is not clear what the point of distinguishing them is, and other moderate realists have identified worlds with sets of propositions.

22However, Aristotle’s theory does not seem to be similar to Plantinga’s in this respect, since Aristotle seems to have thought that states of affairs that are false (i. e. do not obtain) do not exist. Aristotle’s states of affairs seem to be closer in this respect to Meinong’s objectives; Meinong thinks that some objectives (those that are thought of by human beings) are true or false yet false objectives do not exist (but can still have properties and be referred to).
states of affairs would not be the same entities as propositions. However, even if such a development of a representation theory of truth is needful, this would not count against my view that David’s and Crivelli’s states of affairs can be identified with Russellian propositions. David does not distinguish two conceptions of proposition that are often distinguished, Russellian and Fregean propositions. We can say that David’s states of affairs are Russellian propositions, but the propositions that determine them are Fregean propositions.

Many philosophers, e.g. Horwich in [Hor90, §31, page 94], have claimed that we can acknowledge the existence of both Fregean and Russellian propositions. Horwich thinks that Russellian propositions (which he also calls concrete propositions) expressed by a sentence consist of the referents of the words constituting the sentence while Fregean propositions (which he also calls abstract propositions) consist of their senses; Horwich thinks that there are also mixed propositions which contain both referents and senses. David identifies facts with obtaining states of affairs and therefore (since she identified states of affairs with propositions and obtaining with truth) with true propositions, here differing from Crivelli; however, this is a different conception of fact than Armstrong’s.

We can then distinguish from each other four different kinds of purported entities in order of increasing abstractness: facts or states of affairs in Armstrong’s sense, Russellian (or concrete) propositions i.e. states of affairs in Chisholm’s, Plantinga’s and David’s sense, mixed propositions and Fregean (or abstract) propositions. It is quite consistent to hold, so long as we reject mereological extensionality, that all of these entities coexist and are distinct from each other and that the three latter kinds of entities are all truth-bearers. Of course, this does not imply that we would have any positive reason to hold that all of these entities exist, or even that any of them exist (or even that it is metaphysically possible for them to exist); which (if any) of them exist is a further and independent question. However, distinguishing these four conceptions clearly is the first prerequisite for finding any sensible answers.

However, while Crivelli’s interpretation is plausible, there are points on which it can be doubted. Aristotle also says (in Metaphysics 1027b32-33) that what is in the sense of being true is a different thing that is from the things that are in the strict sense (Crivelli’s translation) i.e. that this sense
of being (being qua truth) is different from the proper (κυριωσ, kyrioos) senses (Tredennick’s translation). While this passage is a bit obscure it can naturally be interpreted as saying that states of affairs (which are in the sense of being true or false) do not exist in the strict sense, and this implies that they ought not to belong to our ontology. Furthermore, Aristotle also says immediately afterwards (Metaphysics 1028a1-2) that the cause of what is in the sense of being true is an affection of thought, and this seems to imply that these states of affairs are dependent on human minds, which also suggests an anti-realistic theory of them.

If you want to insist that these things (pragmata) are also truth-makers, you could even find an identity theory of truth in Aristotle as well as a correspondence theory. Aristotle’s theory might then possibly be viewed as a predecessor of the view Russell had in the very first years of the 20th century, but not of the fact-based correspondence theory Russell developed in the 1910s.

While according to Crivelli Aristotle’s theory does not qualify as a correspondence theory on the conception according to which an assertion is

---

Aristotle then apparently says that some items are non-beings in the sense of things that are false. He even thinks that some of these are impossible, since he gives as an example “the diagonal’s being commensurable”. You can then even find an anticipation of Meinong’s theory of objects in Aristotle; P. Thom has done so in [Tho02, page 298], claiming that Aristotle’s ontology is formally parallel with Meinong’s according to which some Objects have Being and some do not (so that neither Being nor Non-Being is pervasive). Indeed, Aristotle’s theory of truth can be interpreted so that it bears surprisingly many similarities to Meinong’s (which would of course be no accident, since Meinong was certainly familiar with Aristotle’s theory to some extent, whether directly or indirectly, and influenced by it). Aristotle’s states of affairs may be in some ways even closer to Meinong’s objectives as well as Russellian propositions. However, Aristotle also says that being is common to everything (e.g. in Metaphysics 1004b; see [Cri04, page 159]), so he may not be consistent in anticipating Meinong if he does so. This contradiction can perhaps also be resolved by taking into account Aristotle’s distinction between different senses of being. One way of interpreting this distinction is to understand Aristotle so that he does think it legitimate to speak about Meinongian impossible objects as a façon de parler, but thinks that these are to reduced away into thoughts as mere logical constructions out of them. This seems to me to be a quite plausible theory when it comes to impossible objects, but less so when it comes to propositions.

---

Even if Aristotle thought of these things (pragmata) as truthmakers these things (pragmata) Aristotle speaks about might also be taken to be more like combinations of tropes (or of tropes and concrete individuals) instead of combinations of individuals and universals. So even if Aristotle hints at a theory of truthmaking in such passages, it may be a theory of truthmaking by tropes rather than a theory of truthmaking by facts. Crivelli considers a similar possible interpretation in [Cri04, page 48] though without connecting it to modern discussions of trope ontologies. In the end Crivelli disagrees with it, but admits that it cannot be ruled out with confidence.
true just in case there is some fact to which it corresponds, yet according to Crivelli [Cri04, page 135] Aristotle’s theory can be regarded as a correspondence theory based on an isomorphism between an assertion and an object which corresponds to the whole assertion. These objects corresponding to the whole assertion can according to [Cri04, page 134] be non-composite items such as esences and incorporeal substances or composite objects, either material substances composed of form and matter or states of affairs composed of universals or a universal and an individuals.

However, there is something strange here; since states of affairs are supposed to be truth-bearers the isomorphism would in some cases be between two truth-bearers. Because of this it apparently does not count as a correspondence theory at all for all assertions in the sense Künnne defines, since the entities to which assertions are related are themselves truth-bearers (though it might count as a correspondence theory for some assertions since essences, incorporeal substances and material substances are presumably not truth-bearers).

If Aristotle’s theory is such as Crivelli says, then though it

---

25 This again raises of the question of how the central concepts in Aristotle’s ontology are to be interpreted in modern terms; you could see these forms of material substances as trope-like entities or as universals. Crivelli notices in [Cri04, page 121] that Aristotle regards the relation of matter to form as analogous to the relation of a material substance to universals. This could be interpreted so that forms are universals. The interpretation of the concept of matter is even more difficult. If forms were interpreted as universals and matter as consisting of bare particulars, then Aristotle’s composite material substances would be similar to facts in Armstrong’s sense, and could even be viewed as truthmakers for statements concerning the existence of material substances and the inherence of forms. While as we have seen it is very doubtful if Aristotle was a realist concerning states of affairs, it is highly likely that he was a realist concerning material substances. Nevertheless, not only is this way of finding a theory of facts as truthmakers in Aristotle a highly speculative interpretation of Aristotle’s theory, but it is not based on Aristotle’s theory of truth but rather on his general metaphysics, so it would in any case not justify us in taking any kind of truthmaker principle to be a part of Aristotle’s correspondence theory of truth.

26 Aristotle’s theory as interpreted by Crivelli is similar to a theory which Marian David calls in [Dav94, page 31] the representation theory of truth, according to which a sentence is true iff it represents a state of affairs that obtains. Crivelli says that Aristotle’s theory is not quite the same as David’s, since Aristotle unlike David allows only affirmative states of affairs, but this is a small difference. David distinguishes this theory in [Dav94, page 33] from the classical correspondence theory, saying that strictly speaking it competes with the classical correspondence theory, but also says that it can be regarded as a revisionist reconstruction of the classical theory. It is rather ironical that if Crivelli’s interpretation of Aristotle is right, then this revisionism leads back to one of the very oldest versions of correspondence theory, namely Aristotle’s theory. Of course, the trouble is in the apparent statement of Aristotle that the states of affairs do not exist in the strict sense, which Marian David does not make - and which neither Chisholm nor Plantinga would make - which probably does not fit together with the representation theory of truth.

27 The same reason naturally makes it dubious whether David’s representation theory
contains elements from the strong correspondence theory, namely the notion of an isomorphism, it may not yet count even as a weak correspondence theory. Besides, there is another reason why it yet may not count as even a weak correspondence theory, since we have seen that Aristotle may have thought that the states of affairs do not exist in the strict or proper sense, and not independently of thought, so a relation of assertions to states of affairs may not count as a relation of them to reality. After all, reality is often (and rather naturally) understood by philosophers, especially ontologists, as the aggregate of all entities which exist in the strict sense (or as some part of this aggregate, such as spatio-temporal entities which exist, as some ontologists understand it) or even as the aggregate of all entities which exist independently of thought. In any this interpretation leaves open the question what kind of theory can be given for the truth of states of affairs (pragmata), since in their case it could not be an isomorphism to an independent object corresponding to them, as Aristotle’s theory as interpreted by Crivelli contains no such object.

Indeed, one might even accuse the theory of containing a circle; the truth of one kind of truth-bearer is reduced to that of another kind of truth-bearer, but the theory yet owes us an account of the truth of these Russellian propositions, i.e. of the obtaining of states of affairs

28David suggests in [Dav94, page 38] that one could hold that the state of affairs that Socrates is dead obtains just in case Socrates instantiates the property of being dead and fails to obtain just in case Socrates does not instantiate the property of being dead. However, rather than saying that sentences correspond to states of affairs that obtain in this sense, one can just as well define truth directly for sentences in this way, saying that the sentence “Socrates is dead” is true just in case Socrates instantiates the property of being dead. The detour through states of affairs appears to be quite useless if we are trying to define sentential truth; it just makes the definition needlessly complex. Indeed, David seems to have later realized this as she has in later writings (e.g. in [Dav09]) paid more attention to object-based correspondence theories which do not make use of states of affairs. I must caution that I am not saying and do not want to say that the postulation of states of affairs would be useless; they have many important possible uses. As David says, they can be taken to be the bearers of modal properties and also the relata of psychological relations like belief and desire. They can also be taken to be not only bearers of modal properties but also to constitute possible worlds as in Plantinga’s theory and hence to provide truth conditions for modal statements. However, the use of them is pointless in the theory of truth for sentences in which modal concepts or concepts of propositional attitudes do not occur.

196
facts, it must be a correspondence of the states of affairs to the objects these states of affairs concern (and perhaps contain), the particulars and universals (or perhaps concrete particulars and tropes). However, in this case one might have immediately defined the truth of assertions or judgments or thoughts with the aid of a correspondence to individuals and universals. Therefore even if Aristotle’s theory of truth is a correspondence theory in the sense in which it is according to Crivelli’s interpretation, it is not a very satisfactory correspondence theory.

Because of these reasons it seems doubtful that Aristotle’s theory can be a correspondence theory unless it is either an identity theory of truth and identity theories of truth count as a limiting case of correspondence theories of truth or it is an object-based correspondence theory as Künne claimed. In any case it cannot be a satisfactory fact-based correspondence theory. However, it is easy to see how one can get a fact-based correspondence theory (whether satisfactory or not) out of Aristotle’s theory with only small modifications, and indeed it seems likely that this is one way in which fact-based correspondence theories have historically arisen in the late 19th and early 20th century. However, most of the ancient and medieval interpreters of Aristotle seem to have either ignored, de-emphasized or explained away his claims about the states of affairs (pragmata), whether as truth-makers or even as primary truth-bearers. For example, Richard Sorabji, who has directed a mammoth project of translating the ancient commentators on Aristotle and therefore studied them quite thoroughly, says in [Sor04, §7(g),

29 There are some exceptions such as the scholastic Gregory of Rimini (and possibly Adam of Wodeham before him, from whom Gregory may have gotten the idea), whose theory of complex significabilia (complex signifables) can be interpreted as a theory of non-linguistic and non-mental propositions as primary bearers of truth and falsity or as a theory of states of affairs as truthmakers. It is not wholly clear which, however; Simons interprets the complex signifiables as truthmakers, but this interpretation is questionable. In order to make any sense of the history we must make clear how the terminology has changed in the course of history; as Norman Kretzmann says in [Kre70], when medievals spoke of a propositio they were speaking not of a propositional content (as is usually done in modern analytical ontology) but of a propositional sign, written or spoken or mental. Gregory said that the complex signifiables were the entities signified by propositions, but here the propositions mean the sentences or judgments that are propositional signs, so when he spoke of propositions this did not commit him to propositions in the modern sense, propositional contents which would be non-linguistic primary truthbearers, but when he speaks of the entities signified by the propositions this may commit him to them. However, it is not clear whether the signification of sentences Gregory spoke of is to be understood as sentences designating states of affairs or having propositions (in the modern and not the medieval sense) as their senses, and therefore it is not clear either whether these complexe significabilia are primary truth-bearers or truth-makers.

197
that according to them the primary bearers of truth and falsity are composite thoughts, not composite sentences nor composite things. Here the composite things clearly correspond to Crivelli’s states of affairs, so that the ancient commentators interpreted Aristotle differently from Crivelli. Because of this their Aristotelian theory of truth was probably in any case an object-based correspondence theory (taking beliefs or judgments as truth-bearers) even if Aristotle’s own theory was not.

In any case when Tarski says that the intuitions to which he wishes to do justice find their expression in the words quoted above, it is likely that he did not want to commit himself to any further aspects of the Aristotelian theory of truth. Clearly there are significant differences between Aristotle’s and Tarski’s theories of truth as well as Tarski’s theories and most theories of Aristotelians. Tarski considers only sentential truth and left open the question of what other truth-bearers there are and what kind of theory might be developed for them, while as we have seen most Aristotelians took the primary truth-bearers to be mental, private entities such as judgments and thoughts and Aristotle himself may have taken either thoughts or states of affairs to be primary truth-bearers. However, this difference is not a contradiction, since Tarski never denied that there were other truth-bearers than sentences nor expressed any opinion about what truth-bearers were primary.

4.2.4 Truthmaker Theories

What are truthmakers?

Tarski’s theory does not imply any truth-maker axiom such as could be used to argue for the existence of facts or tropes. Indeed, we cannot even ask whether the truthmaker axiom is true in the rather poor extensional language in which Tarski’s theory was originally formulated nor in most of the languages in which it is formulated even today, since that axiom cannot be formulated in such an extensional language.

The notion of truthmaking is usually understood in modal terms so that an entity makes a proposition true, i.e. is its truthmaker, only if it is necessary that the proposition be true if the entity exists. For instance, Armstrong says in [Arm97, §8.11, page 115]:

The assumption here is that the truthmaker for a truth must
necessitate that truth. In the useful if theoretically misleading terminology of possible worlds, if a certain truthmaker makes a certain truth true, then there is no alternative world where that truthmaker exists but the truth is a false proposition.

The notion therefore makes use of the modal concept of necessity and must be formulated in an intensional language. If we use the symbol $\models$ for truthmaking, then we can say that $x \models p \rightarrow \Box((\exists y)(y = x) \rightarrow p)$. The implication is sometimes altered into an equivalence and taken as a definition of truthmaking, but most truthmaker theorists think that there is more to truthmaking than just the modal aspect.

**Axiom 1** $x \models p \rightarrow \Box((\exists y)(y = x) \rightarrow p)$.

The principle that every proposition or (interpreted) sentence (of a certain kind) has a truthmaker in this sense, an entity whose existence necessitates the truth of that proposition or sentence, is generally called the truthmaker principle, the truthmaking principle, the truthmaker axiom or even simply the truthmaker. The truthmaker axiom can be formulated as either an axiom schema or as an axiom; the latter involves some kind of propositional quantifiers, whether objectual or substitutional, so it is a more controversial formulation, but it has of course all the advantages a finite axiomatization has. So if $p$ is a propositional variable, the following will give a symbolisation of the standard formulation of the truthmaker axiom:

**Axiom 2** $(\forall p)(p \rightarrow (\exists x)(\Box((\exists y)(y = x)) \rightarrow p))$.

While most truthmaker theorists accept this principle, some truthmaker theorists have rejected this principle or rather this formulation of the truthmaker principle in favour of a weaker principle or formulation. Some of these formulations are rather obscure. Josh Parsons has given a rather clear alternative to the standard truthmaker theory. Parsons calls in [Par99] the view that the standard truthmaker principle holds truthmaker essentialism; other have called it truthmaker necessitarianism. According to

---

30 This expression is not wholly appropriate; there is a sense in which even the (in some ways) weaker theory which Parsons supports is essentialistic. As seen below, he thinks that every true sentence supervenes on the nature of some thing; however, one sense of the word "essence" is identical to the sense of the word "nature" Parsons uses. The expressions "essence of a thing" does indeed usually stand for the conjunction of the
Josh Parsons, the name “truthmaker principle” should be given to a weaker principle according to which every truth has a truthmaker, but this does not have to be essentially a truthmaker of that truth. Parsons would explicate this weaker truthmaker principle so that every true sentence’s truth supervenes on the nature of some thing. He adds that one might gloss a thing’s nature as a grand conjunctive property, conjoining all of a thing’s intrinsic properties. He emphasises that it is no part of the conception of a nature that a thing could not have failed to have the nature it does; so the nature is a conjunction of all intrinsic properties, including contingent intrinsic properties.

The following formula might symbolize Parsons’s formulation of the truthmaker axiom, if $P$ is a predicate variable and $\text{Intr}(P)$ means the $P$ is an intrinsic property.

\[ (\forall p) (p \rightarrow (\exists x)(\exists P)(\text{Intr}(P) \land \Box (P(x) \rightarrow p))). \]

This truthmaker principle must be formulated in modal terms just as much as the standard essentialist principle, since while there are innumerable suggestions as to how the notion of supervenience is to be properly formulated it is generally agreed that it must be formulated in modal terms.

31 Parsons is seeking to defend extreme nominalism and trying to show that there is no valid truthmaker argument against nominalism. However, as he himself admits, this gloss is, to say the least, not obviously compatible with nominalism. Indeed, this is an understatement; we might rather say that it is obviously not compatible with extreme nominalism. It does nothing to help the defence of nominalism to claim that properties are not truthmakers, if the formulation of the truthmaker axiom presupposes the existence of properties in virtue of which individuals are truthmakers. However, while Parsons’s truthmaker theory would not help us to get rid of universals as he himself hoped, it is yet more economical than Armstrong’s, since it does help us to get rid of states of affairs; Parsons’s truthmaker principle only leads to the existence of universals and particulars, which Armstrong also supposes to exist as constituents of states of affairs, but not of the states of affairs Armstrong supposes them to constitute. At least this is so if Armstrong is supposed to be a realist about universals, as he mostly claims to be; at some places Armstrong does seem to suggest that only states of affairs exist, and if this were so, the question of economy would become more complex.
so it also requires an intensional language and cannot be formulated in the language of Tarski’s original theory.

However, even if we admit contrary to Parsons that a truth-maker must necessitate the truth of what it makes true by itself this is apparently not enough, for otherwise every entity would be a truth-maker for a necessary truth, and in this case the truthmakers of a necessary truth can hardly explain its truth. It is a difficult problem how the notion of truth-makers should be characterized more precisely. Many philosophers, such as Gonzalo Rodriguez-Pereyra in [RP02, opage 35], have argued that the notion of truthmakers should be taken as a primitive in the sense that it has no non-circular definition, though it can be usefully clarified in formally circular ways. However, I do not see why taking the notion of a truthmaker as a primitive would be better than just taking the notion of truth itself as primitive, as many philosophers have done, e. g. the early Russell and Moore and later Donald Davidson. Rodriguez-Pereyra says that the lack of a non-circular definition of truthmaking need not be a problem provided we have fairly clear understanding of the concept in question. I am not sure that we have a fairly clear understanding of the concept in question (exactly who are the "we" and how clear would that have to be?). When truthmaker theorists give simple examples to introduce their theory I seem to understand what they are talking about, but when I read their advanced internal debates all the ideas seem to go blurry in my mind; I am practically sure I am not the only one with this experience. However, even if we had a fairly clear understanding of what truthmaking is, the early Russell and Moore could equally well say that the lack of a non-circular definition of truth need not be a problem provided we have a fairly clear understanding of the concept in question. I certainly do not think that we have (at least, I do not have) a clearer intuitive understanding of the notion of truthmaking than of the notion of truth.

Another controversial point with regard to truth-makers is whether we need a truth-maker for every truth. Armstrong defends in [Arm03] Truthmaker maximalism, according to which every truth has a truth-maker (though Armstrong has not always held this view). Kevin Mulligan has also at one time defended this view. However, Barry Smith and Peter Simons have often (e. g. in [Smi99] though here Mulligan joined with them) defended the view that there are truth-makers only for atomic truths.
It seems to me that Smith’s and Simons’s view is rather ad hoc, mainly designed to prevent an argument for the existence of entities they are opposed to such as facts, especially negative and disjunctive facts (i.e., facts that are negations or disjunctions of other possible facts), but also even conjunctive facts that other atomists such as Armstrong accept. Truthmaker maximalism seems to commit us to such complex facts, as it requires truthmakers even for such purely negative sentences as "There are no unicorns.". As Armstrong has shown, truth-maker maximalism commits us to at least one negative fact, which he calls the totality state of affairs. Armstrong does not seem to think this totality state of affairs would be a negative fact, but as it is surely a truth-maker for a purely negative sentence, I would call it negative.

Smith and Simons’s argument why Tarskian account neglects precisely the atomic sentences, though they think it is sufficient in other cases, is very weak indeed. Smith and Simons say in [MSS84, page 288]:

In the wake of Tarski, philosophers and logicians have largely turned their attentions away from the complex and bewildering difficulties of the relations between language and the real world, turning instead to the investigation of more tractable set-theoretic surrogates. . . . we are left with such bloodless pseudo-elucidations as; a monadic predication 'Pa' is true iff a is member of the set which is the extension of 'P'. Whatever their formal advantages, sentences of this kind do nothing to explain how sentences about the real world are made true or false. For the extension of 'P' is simply the set of objects such that, if we replace 'x' in 'Px' by the name of the object in question, we get a true sentence.

The force of this argument is based on claiming implicitly that a Tarskian theory is circular. However, there is surely no need at all to use the notion of truth in characterizing the extension of 'P', at least if 'P' is a genuinely atomic predicate. We can use the notion in doing so, but this does not show the elucidation to be circular; the elucidation would only be circular if there was no way to elucidate the notion of the extension of 'P' without mentioning truth. The notion of extension is just designation in the case of predicates, and the notion of designation need not be explained in terms of truth. The
extension of 'P' is just \{x : Px\}. For instance, if 'P' stands for 'red', then the extension of P can be explained just as the set of all red things. We do not have to say that the extension of 'red' is the set of such things x that 'x is red' is true. We have to make use of the notion of set in the original interpretation of Tarski's theory (though other philosophers such as Field and Devitt have suggested, whether successfully or not, ways of modifying Tarski's definition to get rid of even that notion), and many semanticists are anti-realists about sets who think that sets are just surrogates for real entities. However, not all semanticists think so: an ontologist can interpret the theory of sets realistically and hold that sets exist in the real world and help to make atomic sentences true (though not necessarily in the essentialist sense). Therefore the Tarskian definition of the truth of atomic sentences is not circular and contrary to Simons is a quite genuine elucidation, whether correct or not.

Smith's and Simons's view is also faced with the difficulty of determining or even making sense of the question which truths are atomic truths. We can without difficulty say what are the atomic sentences of a natural language or the atomic formulas of a formal calculus (and hence the atomic sentences of the artificial language we get when we that calculus is given an interpretation). However, a sentence that is atomic in one language can be given exactly the same interpretation as a sentence in another language that is not atomic in that calculus. A sentence of a natural language that is syntactically atomic is very often represented by a complex formula when it is translated into an artificial language based on a formal calculus. Indeed, even two sentences of the same natural language can be necessarily equivalent and probably even synonymous even if one of them is syntactically atomic and the other is not. E.g. "John is bald" and "John has no hair on his head" are surely equivalent and would naturally be said to be synonymous (though some semanticists might disagree) and express the same proposition. Would we say that the statement made by the first has a truthmaker and that made by the second has not? Surely if one of the statements has a truthmaker it also makes the second statement true, as they are necessarily equivalent! Or would we say that the second has a truthmaker just because our language happens to have in it the word "bald"? Surely such a linguistic accident cannot determine what exists in the world, unless we are to be anti-realists such as Simons would not want to be. However, surely
such a linguistic accident cannot even provide any evidence as to what exists in the world. This seems to show that syntactically atomic (structurally identified) sentences of natural languages need not be semantically atomic. Besides an atomic sentence of one natural language can (and even must) often be translated by a non-atomic sentence of another natural language. Thus if a distinction between atomic and non-atomic truths is to have any ontological significance it must be a distinction between different kinds of non-linguistic propositions, not of physically identified linguistic sentences, strings of sounds or marks on paper. Thus truth must primarily be a property of propositions if there is to be an ontologically significant distinction between atomic and non-atomic truths.

However, if truth is primarily a property of propositions, it is doubtful whether propositions can be objectively divided into atomic and non-atomic ones. It need not be the case that every proposition consists ultimately only of atomic propositions. It might be that some positive propositions are endlessly divisible to their conjuncts; e.g. a proposition \( p \) might be a conjunction of propositions \( q_1, \ldots, q_n \), \( q_1 \) again a conjunction of propositions \( r_1, \ldots, r_n \), and so on for ever. It might even be that every true positive proposition concerning particulars would be such an endlessly divisible proposition. I will have already argued that the world might consist of endlessly divisible particulars and the properties of such particulars can also be infinitely complex. If this is so, then the true propositions about them whose structure corresponds to or reflects the structure of the world could also be analogously infinitely complex and endlessly divisible. Even if a sentence of a natural language or even an artificial one is syntactically atomic it may still express an infinitely complex proposition that has no simple constituents. In this case if only simple propositions had truthmakers then if the truth-conditions of complex propositions would ultimately have to be be resolvable to the existence or non-existence of truthmakers for simple propositions it would follow that some propositions that have truth-values would not have any truth conditions at all, which is of course absurd.

Smith and Simons do occasionally ascribe truth to propositions, but do not seem to understand propositions fully realistically, so the ultimate status of the truth-bearers in their theory remains very obscure. Even if truth were primarily a property of sentences, whether there exists a truth-maker for a truth cannot depend on what interpreted calculus or what natural language
the truth is represented in, unless reality itself is to be viewed as relative to
language. Thus the view of Smith and Simons would lead towards an anti-
realism they themselves would not accept, if we would not accept a baseless
fundamentalism.

There is another important problem with non-maximalist truthmaker the-
ories. As Candlish and Damnjanovic see (see [CD07, §3.4]), if all sentences
do not need truthmakers, then a truthmaker theory cannot give a general
theory of truth. Truth itself cannot in that case consist in the existence of a
truthmaker. Nor can such a non-maximalist truthmaker theory be a version
of correspondence theory (as usually interpreted), for correspondence theory
says that truth is correspondence. While sentences with truthmakers could
perhaps correspond to their truthmakers, if some sentences do not have
truthmakers, what would they correspond to? At most a non-maximalist
truthmaker theory could be a version of correspondence theory for some
truth-bearers, not all; however, being a correspondence theorist about some
sentences and not all would obviously lead to an undesirably complex theory
of truth. Even if a truthmaker axiom restricted to some sentences, such as
atomic ones, were true (which I do not want to deny, preferring to suspend
judgement on the matter), it could not by itself answer the question of what
truth is.

Since the truthmaker axiom can only be formulated in an intensional
language, to ask whether the axiom is true Tarski’s theory must therefore
be extended so that it is formulated in an intensional language. Of course
this can be done, and very many semanticists have long since done it. How-
ever, there are no reasons to think that even if we extend Tarski’s theory in
this way, it would have to imply any truthmaker axiom or carry ontological
commitment to facts or tropes. Nevertheless, since such an enrichment of
the language in which Tarski’s theory is formulated would inevitably also
increase the ontological commitments of that theory, such an extension of
that theory might have to carry ontological commitment to properties -
for example if properties are either understood as the intensions of expres-
sions whose extensions are sets or if sets are straightforwardly replaced by
properties as the extensions of predicates - and so be even more blatantly in-
compatible with any kind of nominalism (and therefore also with any kind of
minimalism) than Tarski’s original theory. Therefore Tarski’s theory shows
that a semantical realist does not have to be either a minimalist or a clas-
sical correspondence theorist; there are viable positions between the two extremes. Of course, Tarski’s theory is not incompatible with a truth-maker axiom, as it is incompatible with minimalism, so Tarski’s theory could be combined with a neo-classical correspondence theory though it cannot be combined with minimalism. However, the fact that Tarski’s theory of truth does not require any truth-maker axiom already gives us at least some reason to suspect that no such axiom is needed, since Tarski’s theory seems to answer satisfactorily all the problems connected directly with truth as such. Of course, as we saw that Field has shown, Tarski’s theory leaves open questions about the nature of reference even though reference is central to Tarski’s theory, but theories of facts have surely not been proposed as answers to such questions. Therefore Tarski’s theory gives us reasons to be very suspicious of any classical or neo-classical correspondence theory of truth.

Ilkka Niiniluoto’s Interpretation of Tarski as a Truthmaker Theorist

As David says, a truthmaker theory may be presented as a competitor to the correspondence theory or as a version of the correspondence theory. As if just to make the confusion complete, some recent truthmaker theorists (for example see [Asa11, page 73]) have even said that truthmaker theory is not a theory of truth in the classical sense (since they think it is a purely metaphysical rather than semantical theory) and so is neither a form of correspondence theory nor a competitor to it. However, the notion of truthmaking and the truthmaker principle are even sometimes held to be essential to correspondence theories. Ilkka Niiniluoto says in the abstract of [Nii04, page 57]:

A hallmark of correspondence theories of truth is the principle that sentences are made true by some truth-makers.

However, this is surely historically false. The notion of truthmaking is not a part of traditional object-based correspondence theories of truth. You can scarcely find any explicit notion of truth-making in Aristotle or Aquinas; as I have argued, the states of affairs (pragmata) that Aristotle occasionally speaks about are more likely mere truth-bearers than truth-makers. Modern advocates of truth-makers like Kevin Mulligan, Peter Simons and Barry
Smith in articles such as the classic [MSS84, page 287] refer to 20th century philosophers like Russell, Husserl and Wittgenstein as their predecessors, not to ancient Aristotelians.

Furthermore, the notion of truthmaking is not even essential to all fact-based correspondence theories (nor of course need a truthmaker theory be a fact-based correspondence theory, since there are truthmaker theories such as trope theories which do not make use of facts). While all fact-based correspondence theories say that truth is based on a relation between truth-bearers and facts, yet they may disagree widely about what that relation is. There have been many fact-based correspondence theories which say that true sentences are names of facts or denote or designate facts, and these theories are different from fact-based correspondence theories which say that true sentences are made true by facts. Even Tarski, in the one sentence where he mentions states of affairs, speaks of a sentence designating an existing state of affairs, not of a state of affairs or fact making a sentence true, so even if this one sentence were to be taken at face value, it would not indicate a theory of truthmakers, even though it would indicate an ontology of states of affairs. The relation of designating or naming is very different from the relation of truthmaking: the notion of the relation of truthmaking as it is commonly conceived is a modal notion but the notion of designation or naming is not a modal notion (even if there are interesting connections between naming and necessity as Kripke has shown. There is even a formal difference between designation and truthmaking; a sentence can be made true be many different facts (e.g. sparse truth-making theories which do not admit disjunctive facts hold that a disjunctive sentence is made true by a fact which makes either of its disjuncts true, so if both of the disjuncts are true then both of the facts making them true also make the disjunction true), but

\[\text{32}\] Indeed, the qualification of states of affairs as existing would be redundant unless some sentences designated states of affairs which did not (actually?) exist, so Tarski's statements seems to actually imply a possibilist or even a Meinongian ontology if taken at face value despite Tarski's warnings! It is perhaps no wonder that Künne calls the paragraph where Tarski formulates this irritating; as he says, some fiends of the category of states of affairs find the qualification "existing" less then happy, and tend to say rather that there are states of affairs of which roughly half obtain. Perhaps Tarski should have said "obtaining" rather than "existing" to make his true intention clear. However, this is by no means certain; I argue elsewhere in this work that Aristotle's theory of truth bears many similarities to Meinong's when it is closely examined, so the apparent Meinongian implications of Tarski's formulation might actually fit well with Tarski's stated aim of explicating an Aristotelian conception of truth.

207
a sentence surely cannot be designate many facts; the designation relation is many-one but the truth-making relation is many-many. Of course, it is possible to hold that both relations hold between truthbearers and facts, and some fact ontologists have no doubt done so, but this is by no means necessary and perhaps not even natural. There have also been other relations besides truthmaking and designation suggested as the relations at the basis of truth; even picturing has been suggested.

As the truthmaker principle is then not an essential part of either object-based or fact-based correspondence theories of truth, it is surely not a hallmark of correspondence theories of truth, and therefore Tarski’s theory can be a correspondence theory of truth even if it does not imply or justify the truthmaker principle. Therefore Niiniluoto is correct in claiming that Tarski’s theory of truth is a version of the correspondence theory of truth, but incorrect in claiming that because of this it would be a truthmaker theory.

This, of course, leaves open the question whether Tarski’s theory can be shown to imply the truthmaker principle. Niiniluoto considers a suggestion of P. Kolář in [Kol99] that infinite sequences of objects or sets of them could be truthmakers. However, this theory has fatal flaws. Kolář himself [Kol99, page 77] abandons this idea as a way of interpreting Tarski’s theory as a strong correspondence theory of truth because it would clash with the idea that different truth-bearers are not always made true by the same truth-makers. This is a perfectly valid reason to reject this suggestion. However, there is a a deeper problem with the notion. Even if an infinite sequence satisfies a sentence, yet if the sentence is a contingent sentence attributing an accidental property to an object, the sequence does not satisfy the sentence necessarily. This implies that the sequence can exist but the sentence need not be true, and therefore the sequence is not a truthmaker for the sentence in the most common sense of the word (the sense used by such classical truthmaker theorists as Mulligan, Simons and Smith). The same holds of sets of sequences; any set of sequences can exist and yet a sentence attributing an accidental property to an individual can fail to be true. The basic problem with this kind of suggestion is that notion of truthmaking is a modal, intensional notion and cannot be defined by means of the extensional notions of model theory. Actually, if we want to interpret Tarski’s theory as a fact-based strong correspondence theory in this way, then it might be
better to make use of the version of fact-based strong correspondence theory according to which the relation between truth-bearers and truth-makers is designation, since designation is not an intensional notion; however, the problem Kolář pointed out would still recur, since then all sentences would designate the same fact, and this would surely be unwelcome.

Using sets of sequences and worlds or functions from worlds to sequences (such as are used in the model theory of modal predicate logic) would repair this problem; however, it can be argued that in this case we would arrive at the notion of proposition rather than the notion of fact. Carnap, who generalizes Tarski’s theory in this way (and to whom Niiniluoto refers), does speak of propositions, not of facts or even states of affairs. Such sets or functions would correspond one-to-one with sentences and would function as their meanings (as they do in the semantics of modal predicate logic) and could even be taken to be truth-bearers rather than truth-makers. Facts would then be just true propositions - propositions which contain a sequence containing the actual world - so we would arrive at the identity theory of truth, rather than a fact-based correspondence theory of truth. Of course, the identity theory of truth might also be taken to be version of correspondence theory in a broad sense of the word, so this might not be a bad outcome for a defender of a correspondence theory, but it would make the terminology of truthmaking redundant.

Niiniluoto agrees with Kolář that taking sequences as truthmakers will not work; however, he has a proposal for finding truthmakers in Tarski’s theory of his own which makes use of the formal conditions for truthmaking that Kolář gives. Unfortunately Niiniluoto’s positive proposal is not presented in very much detail. Nevertheless it is clear that there are serious problems with it. Niiniluoto does not analyze the notion of truthmaking sufficiently before setting out to find it in Tarski’s theory, so it is not clear what kind of evidence he would take to be enough for him to prove his claim. However, I will prove that in any case Niiniluoto does not succeed in proving that any standard theory of truthmakers - a theory accepting what Josh Parsons has in [Par99] called truthmaker essentialism - such as the famous theories of Armstrong, Mulligan, Barry Smith, Rodriguez-Pereyra etc. - would follow from Tarski’s theory. Perhaps some non-standard theory of truthmaking could follow from Tarski’s theory, but this would have to demonstrated by defining clearly the notion of truthmaking one claims to follow from it.
Niiniluoto suggests in [Nii04, page 70] that the models of a language can be taken to be truthmakers for the sentences of that language. He goes further and suggests that truthmaking be then identified with the semantic truth relation $|=\$. However, this view is based on a confusion; though the two relations are often symbolized with the symbol, they yet cannot be the same relation. What Niiniluoto apparently means with the semantic truth relation, the notion of truth with respect to a model, is a logical, formal notion, while the general notion of truthmaking is a non-logical and probably material notion. There is much inconclusive debate over exactly what the notion of necessity involved in the truthmaker principle is, but at any rate it is agreed that it cannot be the notion of logical necessity in a narrow sense which model theory can capture. Niiniluoto’s theory of truthmakers does not capture the notion of material necessitation, whether it be some primitive notion of metaphysical necessity or logical necessity in a broad sense which nearly all truthmaker theories agree to be essential to the notion of truthmaking (even if many of them hold that truthmaking cannot be reduced to such necessity). Niiniluoto agrees that Tarski’s theory of truth is a theory of truth for interpreted languages. However, this means that the interpretation of a sentence in such a language is essential to it, and therefore such a sentence cannot be reinterpreted without turning it into a sentence of another interpreted language. This, however, amounts to just changing the subject. Nevertheless, in applying the relation of truth in a model to the sentences of an interpreted language Niiniluoto does reinterpret them, since the notion of a model contains the notion of interpretation, and this means that he does not speak any more of the truth of those sentences but of the truth of sentences of a different language, which means that he just changes the subject.

It is easy to show by means of concrete examples that Niiniluoto’s theory of truthmaking does not work. Niiniluoto’s theory implies that many sentences which are in fact not true are yet made true by existing entities and therefore that they are true. It follows from Niiniluoto’s theory that some models are truthmakers for sentences which (it is pretheoretically obvious) can be false even though a model for them exists and even for some sentences which are false. A model whose domain contains water but not snow and which interprets the word ”snow” as water might make the sentence ”Snow exists.” true according to Niiniluoto’s theory but snow might still fail to
exist even if such a model exists so the sentence might be false. Even more
egregiously, a model whose domain contains water and whose interpretation
function associates the word ”phlogiston” with the word ”snow” would make
the sentence ”Phlogiston exists.” true according to Niiniluoto’s theory even
though phlogiston does not exist so that the sentences is in fact false. At
most the totality of all models would make logically true sentences true but
even it cannot make contingent non-logical sentences true (and cannot even
make materially necessary sentences true) so much less can a single model
do so.

In fact it is doubtful whether the theory of truth to which Niiniluoto’s
theory of truthmaking would lead if consistently followed can even be con-
sidered to be a correspondence theory at all. It has many similarities with
coherence theory, because of which it could also be accounted a coherence
theory, for all that mere model theory can show is that some sentences are
logically consistent and consistency is a kind of coherence. Actually most
historical coherence theorists had a notion of coherence that was stronger
than logical consistency, even though it is usually quite obscure just in what
way it was stronger, so Niiniluoto’s theory would (contrary to his own in-
tentions) be even more radically anti-realistic than most coherence theories.
It would be more appropriate to call the relation of truth with respect to
a model coherence-making (or even better consistency-making) than truth-
making and to call relational structures coherence-makers (or even better
to call them consistency-makers) than truthmakers. In fact if we follow
Niiniluoto in taking models to be truthmakers this leads immediately to
Putnam’s argument against realism. Since Putnam, however, is supposed
to be an opponent of the correspondence theory and of realism while Ninilu-
oto purports to be a defender of both, this convergence of Niiniluoto’s and
Putnam’s ideas shows that something has gone seriously wrong.

Even if the relation of truthmaking cannot be the same as the relation
between sentences and models with respect to which they are true, rela-
tional structures might still be truthmakers, since an additional relation of
truthmaking might hold between then and sentences besides the relation of
truth with respect to a model. Is this the case? We can easily show that
it is not the case, if the notion of truthmaking we are using is the standard
necessitarian notion fulfilling truthmaker essentialism. In [Nii99, page 98]
Niiniluoto suggests that the relational system consisting of John and Mary
and the relation of loving between them is a truth-maker for the sentence
"John loves Mary". However, the problem with this suggestion is familiar.
This kind of theory not fulfil the truthmaker principle as standardly formu-
lated. This relational system can exist even if the sentence is not true, even if
John does not love Mary. Any set theory implies that if John and Mary and
the relation of loving exist then the relational system exists but this can take
place even when John does not love Mary. In fact, this kind of theory does
not even fulfill a weaker principle which Mulligan, Smith and Simons call
(in [MSS84, page 313]) the first principle of truth-making, namely that what
is made true is true. It is then easy to find counter-examples to Niiniluoto’s
theory that show this; let us just take a slight modification of his own ex-
ample. The relational system consisting of Sarah Paley, Barack Obama and
loving exists, since its constituents do, so according to Niiniluoto’s theory
the relational system makes the sentence "Sarah Palin loves Obama." true,
but the sentence is not true, since Palin does not love Obama but hates him.

However, if we use a non-essentialist notion of truthmaking as mere
supervenience of truth on reality such as Josh Parsons proposes, then the
relational structure can perhaps be called a truthmaker; that the relation
contained in the structure relates the other members of the structure might
considered to be a part of the nature of the structure, since it does not
involve relations to any entities outside the structure. However, in this case
the relational structure is already unnecessarily rich; the mere unordered
set or mereological sum composed of the members of the structure would be
enough\textsuperscript{33}. We can even say that the components of the relational structure
are also truthmakers (even if not separately then together) in this weak sense
so the relational structure is in any case not essential to truthmaking even
in this weak and non-standard sense.

Besides, the relational structures are also suited to play another role
in the theory of truth; relational structures can play the role of Russellian

\textsuperscript{33}Indeed, we might even say that in the specific example given by Niniluoto, John
himself might suffice as a truthmaker in Parsons’s sense, since loving Mary can perhaps
be said to be an intrinsic property of John and so part of his nature. After all, people
can love or have other attitudes towards objects that do not exist; John could also love
Galadriel or any fictional object; however, externalist theories of content would object to
the assumption that this is possible in the case of attitudes directed at existing objects.
In any case in the case of other sentences (e. g. "John hits Mary") both John and Mary
are needed for truthmaking in this sense; the mereological sum of John and Mary might
in such cases be the best candidate for a non-essentialist truthmaker.
(though not Fregean) propositions (or what Marian David and Crivelli call states of affairs, though not what Armstrong calls states of affairs) and serve as truth-bearers. If they were also taken to be truth-makers, then Niiniluoto’s theory of truthmaking would lead to an identity theory of truth rather than a classical correspondence theory. This would not show that it is not a correspondence theory, since as we have seen some kinds of identity theories can be taken as versions of correspondence theories, but it shows that it is not a classical correspondence theory as he claimed, so far as we can even legitimately speak of a classical correspondence theory at all. It might be considered as a version of what David calls a representational theory of truth; however, this does not justify attributing a representational theory of truth to Tarski. In any case we have argued that such a theory is either circular or needlessly circuitous.

4.3 Different definitions and kinds of deflationism and inflationism

Perhaps the strongest sense of deflationism about truth (formulated by an opponent of deflationism, Gupta, in [Gup93, 359]) is that according to it the concept of truth cannot play a substantive role in, for example, philosophy of language and metaphysics. However, this characterization of deflationism is somewhat vague or obscure, which leaves its strength also vague; what kind of role is considered to be substantive? The notion of substantiveness at play here seems to be subjective; different people will find different metaphysical commitments substantive, and therefore different people will find different theories of truth to be deflationary. We can distinguish between mild deflationism and strong deflationism, where mild deflationism takes the concept of truth to play a mildly substantive role and strong deflationism which takes truth to have no substantive role at all. Even further, we can distinguish degrees of substantiveness, so that there can be lengthy series of less and more substantive theories of truth. Sometimes - although only very rarely - deflationism is characterized as saying that according to it the theory of truth has no metaphysical implications at all, but this mostly occurs in popular or rather unreliable sources (e. g. Wikipedia’s entry on deflationism has at times claimed this). This would be a rather exact characterization of deflationism; however, it is almost certain that no theory that
has been presented satisfies this definition of deflationism, and certainly no
theory that would be at all plausible. In fact every theory of truth that is at
all plausible has some ontological and therefore metaphysical commitments.
Even the strongest deflationist theories have ontological commitments (e.
ge. as Field admits they always have ontological commitments to the exis-
tence of some kind of truthbearers), and these are often stronger than their
proponents realize.

Tarski’s theory would surely not be a deflationary theory in this strong
sense. Tarski’s theory of truth certainly has ontological commitments ac-
cording to the theory of ontological commitment developed in this disserta-
tion, for example commitments to a variety of set-theoretical entities which
most nominalists would not want to accept into their ontology, such as
the sequences that satisfy sentences and the sets which serve as the ex-
tensions of predicates, since Tarski certainly quantifies over such entities
in the metalanguage. This may be obscured by the perhaps unfortunate
fact that the object language Tarski was using as his example itself con-
tained set-theoretical notions; however, this is not the reason why Tarski
made assertions in his article which carried ontological commitment to set-
theoretical entities. Tarski’s general method of defining truth demands the
use of set-theoretical notions like sequences and sets of sequences in the met-
language even in cases where the object language would not contain any
set-theoretical notions. Because of this we cannot get rid of the commit-
ments of Tarski’s theory to set-theoretical entities just by refusing to deal
with object languages that employ set-theoretical notions or treating object
theories that imply the existence of set-theoretical entities as false. Nor
does it matter that Tarski himself may have been occasionally sympathetic
to nominalism; no matter how reluctant he may have been to endorse the
anti-nominalist implications of his theory, his theory yet has them. There-
fore (since as we have seen in our historical excursus at the beginning of the
dissertation ontology is part of metaphysics or identical with it or a more
fundamental discipline on which metaphysics is founded) if the theory of
this article is correct Tarski’s theory has metaphysical implications. There-
fore a minimalist interpretation of Tarski’s theory is not possible, as is often
thought, nor is Tarski’s theory even wholly neutral with respect to the con-
troversy between philosophical theories of truth, but is flat out incompatible
with a minimalist theory. Of course, this does yet not suffice to rule out an
attempt to replace Tarski’s theory with a minimalist theory nor even some other, weaker deflationist interpretation of Tarski’s theory.

4.3.1 Field’s Two Theories of Truth and Meaning

Hartry Field has been one of the most influential philosophers working lately in the theory of truth. Therefore discussing Field’s views is essential to any contemporary account of truth. However, Field has changed his views radically, and both his earlier and later view have still followers who have developed both of his theories in new directions; e.g. Michael Devitt to a large extent follows Field’s earlier theory. Without doubt Field has contributed many correct and important insights to current semantical discourse. However, I will argue that Field has unfortunately also helped to obscure what the theory of truth is about and what its problems are.

Field’s Earlier Correspondence Theory

Tarski claimed in [Tar44] that he did not make use of any semantical concepts he could not reduce to other concepts. However, it has often been argued that Tarski did not succeed in this aim. Hartry Field argued in [Fie72] that Tarski only succeeded in reducing the notion of truth to other semantic notions. His argument is based on the fact that the definition of designation given by Tarski is list-like; a similar argument was already given by Max Black in [Bla48].

It seems to me that Field is essentially correct. This is one of his important and valid insights. However, there is one important caveat to be made. Tarski surely succeeded in reducing truth to non-semantical concepts, if we accept the extensional conception of reduction according to which just giving a co-extensive complex concept counts as a reduction, as Tarski did. Many of Tarski’s contemporaries thought that this was enough for reduction, and this is especially true of the positivists who had doubted the meaningfulness of the concept of truth. For instance, Carnap said in [Car67, §2, page 6]:

An object (or concept) is said to be reducible to one or more other objects if all statements about it can be transformed into statements about these other objects.

Carnap explained in [Car67, §35, page 60] this with the aid of the concept of coextensiveness of propositional functions. An object a is reducible to
objects b, c, . . . if for each propositional function which is exclusively about objects a, b, c, . . . , there exists a coextensive propositional function exclusively about b, c, . . . . A proposition or propositional function is exclusively about objects a, b, c, . . . , if in its written expression there occur as extralogical symbols only "a", "b", "c", etc. Most of Carnap’s phenomenalistic and behaviouristic reductions were just as list-like as Tarski’s definition of truth. Therefore showing that the concept of truth could be reduced in the way positivists generally tried to reduce concepts was a quite effective ad hominem argument against positivist eliminativists. Many philosophers who carry on the positivistic tradition are still satisfied with this kind of reduction, so Tarski’s attempt at reduction still retains some power as a basis of an ad hominem argument.

However, along with Field, I do not think this kind of reduction would be a philosophically satisfactory kind of reduction. Nevertheless, it is not easy to provide a better theory of reduction, and it is scarcely justified for a philosopher to appeal to the inadequacy of the extensional conception of reduction in objecting to Tarski’s theory if they have no positive suggestion about what a more adequate conception of reduction would involve. Field admitted in [Fie72, page 362] that he does not have a better theory of reduction. However, the discussion over reduction seems to indicate that a better theory of reduction must make use of controversial concepts like modalities or essences; necessary co-extensiveness is surely according to most persons’ intuitions (including the later Carnap, who requires it in a preface to the second edition of Aufbau; see [Car67, page ix]) at least necessary for an intuitive reduction. Field objects that this will not do, for even if there is an intelligible notion of intensional equivalence, his concern is not with analysing the meaning of the word ‘true’ but performing a reduction. This is a very weird objection for someone who was at the time appealing to Kripke’s causal theory of denotation, since Kripke pointed out that all modal i. e. intensional statements are not true by virtue of meaning; even if a necessity is involved in a reduction, it may be an a posteriori necessity

---

34 In a postscript to this article in [Fie01b, page 28] Field says a bit more about what concept of reduction he was employing. He says that he took the concept of reduction somewhat broadly; in particular, he held that it counts as a reduction of a property or relation if you specify the property or relation functionally and show that the property or relation is physically realized. However, this does not make sufficiently clear what Field means by reduction.
such as Kripke defends. Indeed, these a priori necessities for which Kripke is famous are typically cases of reduction; it is necessary a priori that water is \( \text{H}_2\text{O} \) because water can be reduced to \( \text{H}_2\text{O} \), and similarly the reduction of denotation to causal relationships that Kripke considers would also be a case of intensional equivalence. It seems to me that intensional equivalence is necessary for reduction; however, there are reasons to think that even it is not sufficient for it, so this does not clear away all the difficulties. Also these modal concepts are typically seen as just as mysterious as the concept of reference taken as a primitive, and extreme naturalists would usually reject both, so it seems that extreme naturalists who object to the use of the notion of reference as a primitive must accept the extensional definition of reduction if they are to have any definite idea of what a reduction involves, so they cannot have any coherent objection to Tarski’s procedure. Since Field does not know what an adequate reduction would involve, he cannot be sure that Tarski’s reduction of reference is not adequate or that the causal theory of reference he uses is better.

However, Field stresses that he does not mean to suggest that Tarski’s results are trivial. Unlike some others who have made use of his arguments to discredit Tarski, Field himself thought (at least in this article) that Tarski’s results are extremely important, even though they could not provide the kind of eliminativist reduction many philosophers had longed for. Surely any reduction that decreases the number of primitive concepts is an achievement from a purely theoretical point of view; the extreme importance attached to purely physicalistic reductions by many philosophers may derive from its importance to their world view, which is not a purely theoretical matter.

Field refers to a writing of Tarski’s that is not devoted directly to Tarski’s theory of truth where Tarski says that a definition of semantic concepts of truth is necessary to bring semantics into harmony with the postulates of the unity of science and physicalism. Field claims that this shows that one of Tarski’s motives in his theory of truth was to show that semantics was compatible with physicalism, and a reduction of semantic notions would be needed for this cause. However, though Tarski says that physicalism requires the reduction of semantical notions, this does not show even that he himself was a physicalist, much less that his theory was motivated by physicalism. As Greg Frost-Arnold argues in [FA04], Tarski is likely to just have been attempting to appease an audience of physicalists that he viewed as hostile.
to his ideas, namely the members of the Vienna Circle, by appealing to their views which he need not himself have shared.

In any case, even if this was one of Tarski’s motives it can surely not have been the only one; Tarski said that he was trying to do justice to the intuitions which adhere to the classical Aristotelian conception of truth, and while it is controversial just what such a conception is, yet for any plausible interpretation of that conception, such a conception has been held by many philosophers who have not been physicalists (even if most of them have been naturalists in some weak sense). Surely the first philosopher who has generally been held to be a correspondence theorist, Aristotle, who believed in immaterial Intelligences, the Prime Movers, or still more later Aristotelians like Aquinas, who believed even human souls to be capable of existing after the destruction of the body, were not physicalists. Nor were later correspondence theorists always physicalists; phenomenologists like Husserl and Reinach were surely not, and even Russell, who is usually held as the paradigmatic classical correspondence theorist, was at first some kind of dualist and then a neutral monist rather than a physicalist. Many historians of philosophy (e. g. Peter Simons, Jan Wolenski and Artur Rojszczak - see [Roj02]) have claimed that the philosophical background of Tarski’s definition is not so much the Vienna Circle but rather the tradition of the Lvov-Warsaw School, which ultimately goes back to Brentano (though Brentano’s evidence theory of truth was definitely not a correspondence theory, so this is doubtful) and to Bolzano (who clearly was a correspondence theorist, so this is more plausible). While some philosophers in this tradition like Tarski’s immediate teacher Tadeusz Kotarbiński were indeed physicalists, the earlier philosophers taking part in the tradition were not; Brentano was a dualist while Bolzano was a Leibnizian panpsychist. Tarski’s theory was in any case not compatible with Kotarbinski’s philosophy, since it made extensive use of sets whose existence Kotarbinski denied. Tarski could have succeeded in explicating the classical correspondence theory of truth even if his theory implied that it was not compatible with physicalism, though of course it at worst leaves the question of the compatibility wide open.

Field tries (see [Fie72, page 367]) to develop a definition of denotation and then of truth that would be compatible with physicalism by making use of Saul Kripke’s causal theory of denotation. However, it is highly doubtful if Kripke’s theory is suitable for his purpose. Kripke himself is
not a physicalist, and indeed at the very end of the same work where he
developed his causal theory of denotation, as the culmination of his work,
he argued against physicalism (see [Kri72, pages 334-342]) on the basis of
his theory of denotation. Kripke did indeed argue directly only against a
specific version of physicalism - the Identity Theory - but it seems clear that
he thought that his argument counted against physicalism as such. Therefore
Kripke would not have been interested in developing a theory of denotation
that was compatible with physicalism, so it would be strange if he had by
pure accident developed such a theory. Also Kripke himself did not apply his
causal theory of denotation to all terms but only to proper names and natural
kind terms; he admitted that there were many terms of which the descriptive
theory of truth was correct, and "truth" is plausibly such a term. At least
Kripke’s theory would then need lots of modifications and supplementations
before it could be yield a physicalistic theory of denotation.

In any case, the causal theory would have to be developed in far greater
detail than it has been in order to show that it worked and was compatible
with physicalism (even if we had an adequate theory of reduction and knew
what counted as a physicalistic reduction of semantical notions; since we
do not even have that the prospects of showing that a causal theory of
denotation was compatible with physicalism are dim indeed!). All that the
causal theory of reference says is that there must be some sort of causal
chain between an expression and the entities it refers to; this is very plausible
indeed, but very vague, as it does not say much about what kind of causal
chain is in question. The causal theory as it stands now is more a manifesto
for a research program to be initiated than a finished theory. It is quite
likely that in order to specify the causal chain in question we would have to
appeal to non-physicalistic notions, if we would be at all capable of specifying
it. Nevertheless, I do not want to commit myself to the statement that
the causal theory of denotation would be false; the questions are far too
difficult to give any firm answers to in the current state of philosophical
research, and again, the theory would have to be specified far more precisely
even in order to be shown wrong even if it were wrong. My theory of
ontological commitment and truth will be fully compatible with a causal
theory of denotation as well as almost any realistic (or as it is sometimes
called referential or representational) theory of denotation.

In any case, my motive for using Tarski’s theory of truth in this work is
not to defend physicalism, but to defend semantical realism and show (here agreeing with the early Field) that it does not require a truthmaker principle or fact ontology but does imply significant ontological commitments. Even most physicalists agree that realism is a weaker theory than physicalism, so the reduction of semantical notions is not necessary for my purpose, but taking them as primitive is legitimate for my purpose.

Field’s later Deflationism

Field proclaimed his conversion from correspondence theory to deflationism in [Fie94] (reprinted in [Fie01b, pages 104-140]). Since Field is one of the most influential philosophers writing about the theory of truth, this article has been very influential in spreading deflationism. However, while Field is usually a very clear writer, this article is remarkably obscure in many ways, which may not be apparent on a casual skimming of it; it leaves it rather unclear just what the kind of deflationism he converted to was and just why he converted to it, and therefore whether he was at all justified in doing so. In fact, these reasons, which I will proceed to specify in detail, make it likely in my view that Field’s earlier theory, sketchy and promissory though it was, was better than his later one.

Field begins this article very obliquely by distinguishing two traditions in the philosophy of language and the philosophy of mind. According to him one of the traditions, whose early advocates include Frege, Russell, the early Wittgenstein and Ramsay, says truth conditions play an extremely central role in semantics and the theory of mind; a theory of mind is in large part a theory of truth conditions. As to the other tradition, Field says only that the verifiability theory of meaning provides a crude paradigm of it. He tells us that in this the main concept are not truth conditions but verification conditions. Field sketches this theory so that the verification conditions of a type of utterance might be given by the class of sensory stimulations that would or should lead to the acceptance of an utterance in that class. Field does not even give any examples of representatives of this tradition as he gave of the other tradition, so it can be doubted whether this supposed

---

35 Ayer would seem at first sight to be the best possible representative of this tradition, being both a deflationist (as seen in Chapter 5 of [Aye36a]) and a verificationist. However, there is the big problem that rather than replacing truth conditions with verification conditions in the theory of meaning, Ayer identified the two kinds of conditions. E.g. Ayer said in [Aye34, page 337]:
second tradition has even existed, and it can be suggested that it is just a
research program newly cooked up by Field. Nor does he even tell us whether
there are more than two traditions or whether all theories of meaning and
content should fall into one of the two traditions.

Even more importantly, Field never gives any reasons for thinking that
the second tradition or research program is better than the first. There are
of course many arguments in the literature for such a tradition (as we have
already in part seen in this dissertation), and Field may just suppose that
the reader is familiar with them, though it is far from clear which of these
arguments Field himself would accept. However, there are also well-known
difficulties with this tradition. In fact because of the well-known difficulties
with verificationism, which I went through in Section 3.3 of this dissertation
(and which Field does not address in the article, or anywhere as far as I
know), I would think the first tradition definitely better than the second,
if they did indeed exhaust the options in the general theory of meaning.
However, it is by no means clear that they do. In fact there are reasons to be
dissatisfied with both traditions; meanings and content are naturally viewed
as more fine-grained than either truth conditions or verification conditions
(as indeed Field himself had recognized in other articles such as the one
collected in [Fie01a]). Therefore it may be possible and advisable to reject
both traditions.

This characterization of deflationism is more narrow than Gupta’s, since
Field speaks only of a role in the theory of meaning and content, not in
philosophy of language and metaphysics generally, but it shares the same

Ayer then thought contrary to Field that truth-conditions played a central role in seman-
tics, but verification conditions played an equally central role according to him since they
played the very same role. As we see from the quotation above, he slid instantly and un-
selfconsciously from truth conditions to verification conditions, seeing no difference at all
between them. In this he was like most of the early positivists. Carnap perhaps saw the
ostensible difference between the two kinds of conditions but explicitly identified them,
as seen e. g. in [Car32b, pages 221-222]. Therefore even a logical positivist like Ayer
or Carnap is not a good example of Field’s supposed second tradition. The same holds
of at least many other current verificationists. Michael Dummett did initially suggest a
theory like the one Field describes, and Dummett may have influenced Field (though two
like-minded thinkers scarcely constitute a tradition). However, Dummett later rejected
this suggestion as misleading and also explicitly identified conditions of correct assertion
and truth-conditions in [Dum78, page xxii].
kind of vagueness. The notion of centrality is as vague as the notion of substantiveness. What is a central role? Centrality is intuitively a matter of degree; truth conditions can play a more or less central role, and so we can again distinguish between milder and stronger forms of deflationism, indeed a continuum of them. Field indeed admits that the notion is not initially clear, but promises that his goal is to clarify it as he goes along: unfortunately, this is not always a very good way of making one’s notions clear, and at least as far as I am concerned, the notion never gets very clear in this case.

More importantly, deflationism as Field or Gupta define it does not seem to directly concern the theory of truth in the usual or at least traditional meaning of the word, i. e. attempts to solve the question of what truth is, at all. Field himself argues primarily for (see [Fie94, page 252] and [Fie01b, page 107]) a deflationary view of meaning and content or a deflationist view about the role of truth conditions in content, not for a deflationary view or theory of truth. Surely questions of how truth is to be defined are different from questions about how the notion of truth is to be used. Of course, what a philosopher thinks truth is affects what he thinks it can be used for, but it does not determine it completely; whether truth plays a central role in the theory of meaning depends not only on what truth is but also on what the correct theory of meaning is like. However, Field also mentions (in [Fie94, page 250] and in [Fie01b, page 105]) a deflationary conception of truth, which is in his view one way in which a verificationist can speak about truth and various disquotational theories, starting from a pure disquotational theory.

36 Some of Field’s formulations and arguments actually seem to suggest a very radical view, according to which truth conditions could not play any role. E. g. Field says on [Fie94, page 272] and on [Fie01b, page 128] that the deflationist/inflationist contrast was explained in terms of whether truth conditions play a role in semantics and the theory of content. Here Field does not use the word ”central” at all and seems to deny that truth conditions could play any role in such a theory. If deflationism about meaning and contents boiled down to this claim it would be a quite clear and sharp if implausible claim. However, it would be eliminativist rather than deflationist. Dummett also came close to eliminativism when he said in [Dum78, page 11] that we should abandon the notions of truth and falsity altogether, but he denounced this view afterwards. This kind of claim could be justified by arguing that physicalism implies that sentences could have no truth conditions at all. However, Field says much that seems to be inconsistent with this claim. He does say that a deflationist can speak about truth conditions. However, as I will show in a footnote below, there are reasons to think that his attempt to show how a deflationist can speak about truth conditions does not work, so there are reasons to suspect that Field’s premises might consistently followed out lead to eliminativism rather than deflationism.
as examples of this deflationary theory of content\textsuperscript{37}.

However, what Field says leaves it quite obscure whether there are any ways to implement a deflationary theory of content other than verificationism and the deflationary conception of truth and if so what they might be and why the deflationary conception of truth would be better than them. It also leaves it obscure whether there might be other reasons to accept a deflationary conception of truth than verificationism and whether one might consistently accept a deflationary theory of truth without accepting a deflationary theory of content generally. We must recall that of the logical positivists Schlick had combined a correspondence theory of truth with a verificationistic theory of meaning in \textsuperscript{[Sch79b] and in [Sch34]} (translated in \textsuperscript{[Sch59]}), and if this combination is consistent then the acceptance of verificationism would not force one to reject correspondence theory as the later Field does. However, Field does not try to show that such an option would be impossible. It does seem possible for a verificationist to claim that truth consists in a correspondence between truth-bearers and observable entities, whether macroscopic physical entities or sense-data. Field then does not give good reasons either to accept verificationism or to proceed from such an acceptance to the rejecting of correspondence theory of truth, as he

\textsuperscript{37}Field’s definition of disquotational theories, however, is peculiar in one way. Field has given statements regarding the kind of equivalence that occurs in the T-schema or Convention T that do not seem to fit together very well. Field actually says in \textsuperscript{[Fie92, page 322]} that the equivalence should be logically necessary. Field describes this equivalence in \textsuperscript{[Fie01b, page 114]} as conceptually necessary, which need not be quite the same thing. However, in \textsuperscript{[Fie94, page 250]} and in \textsuperscript{[Fie01b, page 105]} Field tells us that the pure disquotational theory says that "p" and "it is true that p" are cognitively equivalent. The problem is that it is not clear that cognitive equivalence would imply logical equivalence or conceptual equivalence or even material equivalence, and the last at least is clearly required for Convention T or T-schema as they are commonly understood. Field says in \textsuperscript{[Fie01b, page 106, footnote 2]} that his preferred reading of cognitive equivalence is that to call two sentences cognitively equivalent for a person is to say that the person’s inferential procedures license a fairly direct inference from a sentence containing an occurrence of one to the corresponding sentence with an occurrence of the other substituted for it. It is not wholly clear to me what Field means with an inferential procedure, but the most natural way of understanding that notion is a purely psychological one, according to which an inferential procedure licensing an inference from on sentence to another is just a disposition to change (under some unspecified triggering conditions) from a state in which one believes one sentence to a state in which one believes also the other. Now it seems to me that if a person is irrational then the inferential procedures in that sense of that person can befallacious so that p and q can be cognitively equivalent for him even when it is not the case that p if and only if q. This, of course, implies that Field’s attempt to reconstruct the notion of truth conditions by means of a pure disquotational notion of truth fails.

223
does. So he surely does not give any good reasons for his rejection of the correspondence theory of truth.

Nor is it clear whether there are any varieties of the deflationary conception of truth besides the different kinds of disquotational theories Field mentions and if so what they might be and what might be common to all of them. Furthermore, the two traditions Field distinguishes are not connected with specific theories of truth, and therefore by bringing them into the discussion of truth theories Field has helped to sow confusion about what he theory of truth concerns. The first tradition includes according to him both Russell, who was a paradigmatic correspondence theorist (and surely an inflationary correspondence theorist, as his version of correspondence theory was based on a fact ontology), and philosophers who were not correspondence theorists at all. Frege, whom Field presents as one of the founders of the first tradition, notoriously did not hold a correspondence theory of truth, though the second of the founders of the tradition, Russell, did. Indeed, one of the reasons why Frege apparently rejected correspondence theory was that in his view truth was too fundamental to be defined (see [Fre56, page 291]); so Frege rejected correspondence theory (so far as he did) just because truth had a too central role in his view. Frege himself has sometimes been counted as a deflationist, though he is probably better counted as a primitivist about truth. However, the first tradition includes also Ramsay, who is generally regarded as the first deflationist and the founder of deflationism, and who is therefore a paradigmatic deflationist.

The traditions or research programs Field postulates do not therefore seem to be associated with definite theories of truth at all. Nevertheless, Field presents a deflationary theory of meaning and content as incompatible with the correspondence theory of truth, and most of today’s philosophers concerned with truth have followed him in this. It is certainly not at all obvious that it should be incompatible with it. In fact it seems to me that Field’s deflationism about content is probably indeed incompatible with the specific version of correspondence theory he himself had earlier held, based on Tarski’s theory but modifying it considerably, and with some other modern versions of correspondence theory. However, there is no good reason

---

38Davidson advocated most prominently the view that a truth-conditional theory of meaning can be based on Tarski’s theory of truth, and many Davisonians have followed their master in this. However, Davidson also eventually came to the view that Tarski’s theory (as he interpreted it) was not a correspondence theory, so this does not give very
to think that it would be incompatible with all versions of correspondence theory or with (all interpretations of) Tarski’s original theory sans Field’s modifications of it. Deflationism about content need not then be incompatible with the correspondence theory of truth as such.

Field’s sketch of the traditions leaves then very obscure both what their limits are and how they are related to theories of truth. Recently, there has been much debate about the relations between theories of truth and meaning. One vital question is whether deflationism concerning truth is consistent with truth-conditional theories of meaning. There is an influential argument by Dummett and Davidson that these two theories are not consistent, as combining them would lead to vicious circularity; however, recently many philosophers such as Mark Lance (in [Lan97]), Michael Williams, Max Köbel and Alexis Burgess (in [Bur11]) have argued that they are consistent. There is then no agreement at all on this question, so the answer to it cannot be obvious as Field assumed. In any case, even if deflationism were inconsistent with a truth-conditional theory of meaning, this would not show that it was inconsistent with correspondence theory of truth, since as I have shown the correspondence theory does not imply truth-conditional semantics.

However, the pure disquotational theory Field formulates contains some very specific and strong (and intuitively highly implausible) restrictions on a deflationary theory of truth - such as the radical claim that a person can meaningfully apply a pure disquotationalist truth predicate only to sentences he himself understands, i.e., sentences in his idiolect. This surely makes a purely disquotational truth predicate intuitively very different from the strong support to the view that truth-conditional semantics and correspondence theory of truth should be connected.

39Burgess’s claim is weaker than that of the others in an interesting way he does not himself notice. Burgess defends the compatibility of deflationism with mainstream semantics, which he describes both as truth-conditional and model-theoretic and as formal and linguistic. However, as I have already explained in Section 4.2.4 of this work, there is a difference between truth-conditional semantics and model-theoretic semantics. As two interpreted sentences could be true in the same models but yet have different truth-conditions, truth-conditional semantics goes beyond model-theoretic semantics. Unlike model-theoretic semantics, truth-conditional semantics tries to account not only for formal but also for material meaning. Therefore some kinds of deflationism might be compatible with model-theoretic semantics but yet incompatible with truth-conditional semantics. Burgess admits that there is a kind of circularity in combining model-theoretic semantics with deflationism but denies that it is vicious. It is notoriously difficult to decide what kinds of circularity are vicious, so the question is far from simple.
ordinary truth predicate. Many philosophers who call themselves deflationists would not agree with these restrictions. E.g. Horwich says in [Hor90, page 106] that legitimate instantiation of the disquotational schema cannot be restricted to sentences in our current language, but we must be allowed to instantiate an unspecifiable, unlimited number of future sentences of our language. Field himself in a review of Horwich’s book [Fie92, page 325] says that he takes it as a corollary that we do not now understand those instances. Horwich’s theory is then not deflationist in the very strong sense in which a purely disquotational theory would be, and so is not really minimal as Horwich calls it, as a pure disquotational theory would be weaker than it.

I would certainly want to defend a more inflationary theory of both truth and ontological commitment than a purely disquotational theory of either would be. Nor do I see that Field gives any good reasons for accepting or even considering seriously a purely disquotational theory, for it surely would not follow even from any reasonable version of verificationism. A verificationist would say that we can only apply the concept of truth to sentences speaking about observable entities. However, surely for any person there are lots of sentences in other languages which he cannot understand which yet speak about observable entities, even entities observable to him. A verificationist can then apply the concept of truth to such sentences, but a pure disquotationalist cannot. However, as the main reason Field gave for considering deflationist theories was based on verificationism, he surely then gave no reasons for considering purely disquotational theories seriously. Also applying the truth predicate to sentences we do not understand does not imply that we would have to take truth as causally explanatory as Field’s previous theory did, for a property applying to sentences we do not understand might still be an epiphenomenal property.

Nevertheless, there still remain problems even in distinguishing Field’s version of deflationism from inflationary correspondence theories. Though Field formulated the pure disquotational theory, Field has never said outright that he would accept the pure disquotational theory, probably sensing that it is absurd, and even if Field might have at one time accepted that theory, he has clearly come to reject it. The problem is that when Field later in his article (in [Fie94, §8-11]) and even more in the postscript in [Fie01b, pages 141-156], tries to defend deflationism against the obvious objection
that it is just absurd, he seems to modify and weaken its deflationary nature. While a pure disquotationalist theory of truth is clearly different from an inflationary correspondence theory, Field’s quasi-disquotational truth and extended disquotational truth are already hard to distinguish from inflationary truth concepts. Furthermore, it is far from clear that they are any longer compatible with a deflationary theory of meaning and content. Field indeed admits in [Fie01b, page 119] the possibility that as the deflationist tries to explain various facts, he will have reconstructed the inflationist’s relation ’S has the truth-conditions p’. Just as logical positivists discovered that the weakening of verificationism, with which Field’s deflationism is connected, threatened to make it vacuous, so analogously and hence predictably does this weakening of deflationism threaten to make it vacuous in the sense that it becomes indistinguishable from a weak correspondence theory. Field goes on to weaken these modified forms of disquotationalism even further in section 4 of his postscript, where he apparently gives up the main point in which his theory clearly differed Horwich’s and about which he criticized Horwich in [Fie92] as we have seen. Namely, he here seems to admit that an agent can apply a disquotational truth predicate to utterances he himself does not understand.

Field explains that all he really hopes to motivate in his article is that we should be methodological deflationists, i.e. assume deflationism as a working hypothesis. However, I cannot see how he could possibly hope to motivate it, as he nowhere motivates the adoption of the tradition including verificationism that is supposed to motivate deflationism (or even shows that such a tradition exists) or even defends this tradition against the well-known objections to the only apparent candidates for membership in that tradition. Field does not even make it clear to the reader why he himself converted from the inflationary correspondence theory he held earlier to methodological deflationism, much less provide any good reasons for the reader to follow him. All he does is sketch what the research program based on methodological deflationism would look like, but this gives no reason to prefer it to the research program based on a causal theory of reference, which is equally sketchy but no more so.

Field’s theory seems to be threatened with all the problems that we have become familiar with in the history of logical positivism. There are, however, differences between Field’s theory and that of logical positivists.
We must examine whether these are sufficient to save Field’s later theory, i.e. whether his deflationist theory of meaning and content could motivate a deflationist, specifically disquotational theory of truth independently of any pre-existing tradition, which we have already seen either does not exist or is a failure and hence cannot successfully motivate any theory of truth. I will argue at this last part of this subsection of my dissertation that there are reasons to think they are not, and this will suffice to refute methodological deflationism, though of course it does not prove that no deflationary theory of content could ever be made to work.

The main difference between the version of verificationism Field sketches and most earlier versions of verificationism is that where logical positivists such as Ayer originally took the verifying experience to consist of sense-contents, Field speaks of sensory stimulations, which gives his theory a more physicalistic veneer. Of course, here Field is just in line with the way logical positivists themselves, following Neurath, later modified their theory. Field’s formulation is specifically similar to the way the later Quine, in whose theory ‘stimulus meaning’ is a central concept, developed verificationism. However, this attempt at amalgamating verificationism and physicalism shares all the well-known problems of verificationism and physicalism and causes novel difficulties of its own. The use of the notion of stimulation suggests a behaviourist theory of the mind, though Field himself has in other writings (e.g. in [Fie01a, pages 56-58, 62, 74-75]) suggested a functionalist theory. There are well-known arguments that the two theories Neurath and Quine attempt to combine are incompatible, which I have already gone through, such as the possible ambiguity of the word “stimulation”, and many of these apply also to Field’s sketch of verificationism. Also Field’s use of the normative term “should” raises questions about how to understand it; Field’s

\footnote{I distinguished three senses of the word “stimulus” in the case of Quine: the perceived object or event, the sensory intake and the neural input in the brain. It is not at all clear which of these Field intends to denote with his use of the word “stimulation”, and every option presents problems of its own. As we are usually not aware of the neural input (at least as such), it cannot verify any statements as a neural process. However, neither can the other kinds of stimulations verify anything by themselves, as they can occur without any perception or even without bringing about any reactions, if the connection to neural input happens to be absent. However, an experience which verifies statements might supervene upon the neural process or be a functional role it plays (as Field’s own statements in [Fie01a, pages 56-58, 62, 74-75] would suggest) or be an aspect of it or be related to it in some of these ways, so it would seem that some modification like this would be the next step in turning Field’s formulation into a less crude one.}
physicalism, as a form of naturalism, can hardly accept primitive normative properties any more than primitive semantic properties, but it is far from clear how Field would reduce this normativity to anything physicalistically acceptable.

Field’s sketch of verificationism sounds like it presupposes a strong concept of verification, conclusive verification. If all that were demanded were confirmation, then a single class of sensory stimulations could hardly suffice to determine the meaning of an utterance type, as Field’s sketch demands, since any statement can be confirmed in innumerable ways by all kinds of different evidence. As successors to logical positivism such as Hempel and Quine have shown, if assertibility conditions were confirmation conditions, then only whole theories could have assertibility conditions, not single utterances (if even whole theories could have definite verification conditions)\(^{41}\).

\(^{41}\)On the other hand, Field himself had in other writings favoured a conceptual role semantics, which is generally associated with holism, and even in the article in question he hints that verificationism should be combined with conceptual or computational role semantics. Therefore it may be that Field’s sketch of verificationism sounds atomistic just because it is, as he admits, a crude paradigm of the tradition he defends. However, Field only suggests supplementing verificationism in the theory of meaning with other factors like indication relations, not modifying his formulation of verificationism itself. Surely this is not enough to remove the defects of his formulation. Even if Field is not committed to his crude formulation of verificationism, yet he has said nothing about how his deflationism would fit with a more subtle holistic verificationism he really believes in, and in such circumstances it is hard to evaluate his theory at all. The Kripkean causal theory of reference Field had championed earlier was of course an atomistic or localistic theory, so if Field had now switched to holism this was another big change in his thinking he did not make clear to the reader. Surely Field would have to provide more details before his theory could be considered justified in any way. Also a holistic theory of meaning causes difficulties about the possibility of communication, as we have already seen thinkers like Fodor and Lepore have argued. It seems such difficulties also afflict Field’s theory. Field says explicitly in the preface to [Fie01b, page viii] that the deflationary picture gives a strongly first-person orientation to the theory of truth and reference. Such a first-person orientation is what is often also known as the phenomenological perspective. However, such a perspective drove the founder of phenomenology Husserl to transcendental idealism, and it is not at all clear how such a first-person orientation fits together with Field’s rather strong physicalism. A pure disquotational concept of truth only applies to sentences in a person’s idiolect, a private language. Other philosophers who are like Field both physicalists and deflationists like Quine have argued (e. g. in [Qui68, page 185]) that there cannot be such a private language, so Field’s position is rather weird in this respect. Can Field avoid the conclusion that the same sentence has entirely different contents for different people? The same physical entity can cause entirely different sensory stimulations to different people whose sense-organs are different. Do not these people then associate entirely different verification conditions with the same sentence? Also indication relations also seem to vary widely from person to person. If an atomic physicist makes a claim about electrons, this makes it far more likely that the electrons are as he says they are than if an ordinary person makes such a claim. Field states that indication relations are supposed to give meaning a social aspect. Unfortunately he does not go into any details,
However, if meaning were to consist in methods of conclusive verification, then statements referring to theoretical entities (such as subatomic particles) would be meaningless, since they are not conclusively verifiable (or have all the same meaning, the same null conditions).\footnote{It seems unlikely that Field would be willing to accept the kind of radical anti-realism to which there are good reasons to think verificationism leads. Field has indeed been very hospitable to anti-realistic theories of mind (though he shrinks from full-fledged eliminativism) and meaning and universals. There are indeed some signs that he may have relented lately of his opposition to universals, considering in [Fie98] seriously a theory he calls plenitudinous platonism - apparently the same kind of theory propounded by Mark Balaguer in [Bal98] - which might be a form of realism about universals. However, since Balaguer himself says (in [Bal98, page 152]) that this kind of theory is similar to fictionalism, it seems that it is not genuinely realistic, but just a more subtle kind of anti-realism. However, Field seems in any case quite unwilling to accept anti-realism concerning physical objects. Even more Field has often (e. g. in [Fie80, page 34]) shown himself hospitable to substantivalism concerning space-time, a theory which is a notorious example of an unverifiable theory. Indeed, Field’s realism concerning space-time forms an essential premise in his argument against realism concerning universals in[Fie80]. This separates him sharply from the positivistic tradition where verificationism originates, to which Mach’s rejection of substantivalism was crucial. Substantival space-time is surely a theoretical entity if any, and it is very unclear indeed how statements concerning substantial space-time could be given meaning by means of any stimuli! Perhaps Field has rejected this substantivalism when embracing verificationism; however, I do not find that he anywhere says that he has done so, and therefore there exists the danger of a radical inconsistency in Field’s theory as a whole. Field’s introduction of indication relations, which I will discuss below, might help this problem, but introduces other problems of its own.}

It is not clear whether Field’s addition in [Fie94, pages 253-256] (reprinted in [Fie01b, pages 108-112]) of conceptual or computational roles and indication relations to assertion conditions as components of content differentiates his theory enough from the theory of logical positivists to remove this inconsistency. The most promising part of this is the addition of indication relations, which Field says are intended to give content externalist and social aspects. This does seem to indeed go beyond what even the most flexible positivism would allow, since Field allows the relata of this indication relation to be unobservable. However, the introduction of indication relations does not seem to help much with the construction of content or meaning. Belief states that would intuitively be said to have the same meaning can have entirely different indicative relations. Intuitively we would say that the same belief can be held both irrationally and rationally, but a belief has which makes his claim quite unconvincing. All he says in [Fie94, page 256] is that

the conceptual roles and indication relations of other people’s states of believing a certain sentence are counted as relevant to the content of my state of believing that sentence

Exactly how are they counted as relevant to that content?
indication relations to the state of affairs it would be said to be about only when it is rational. Indication relations seem to be better suited to account for justification in an externalist epistemology, as in Goldman’s reliabilism (see [Gol86]), than for content. Indication relations cannot then account for the content of irrational beliefs, at least not alone. Field seems then to confuse epistemological and semantic concepts just as much as logical positivist did, even with his addition of indication relations, though that addition changes the way they are confused.

On the other hand, the verification conditions contained in the conceptual role would do a better job in the case of many irrational beliefs, namely those that concern observable entities. The conceptual role, however, cannot give meaning to sentences speaking about substantival space-time, or even most physical objects, which Field would want to be meaningful, though indication relations possibly might. So it seems that neither indication relations alone nor conceptual role alone can serve to reconstruct content.

However, Field leaves it quite unclear how these three aspects of content are to be integrated. He does say that the conceptual role of a belief state includes its verification conditions but says nothing about how conceptual role and indication relations are to be combined. This is is big defect because this combination is quite problematic, as the two ingredients often when they both exist so to say pull in different directions. It is quite possible if a person is sufficiently irrational or just sufficiently misinformed that the belief state of that person would be verified by a class of sensory stimulations whose absence it indicates for that person. Conceptual roles and indication relations seem combine as well as fire and water.

Furthermore, it is part of our common notion of meaning and content that a person can have untrue and irrational beliefs whose content concerns unobservable entities (and can meaningfully speak about such entities with-

\footnote{Field himself gives a counterexample to trying to construct truth conditions from indication relations (even though he also says he doubts that there is a clear matter of right or wrong here). This counterexample (which in my view is quite clearly a counterexample) also serves as a counterexample to an attempt to construct meaning or content from indication relations. Field considers in [Fie01b, page 110] an ancient Greek whose judgements that Zeus is throwing thunderbolts are correlated reliably with (i. e. indicate) the presence of lightning in his vicinity. Field admits that we would not say that the truth-conditions of his utterance “Zeus is throwing thunderbolts” are just that there is lightning in his vicinity. However, neither is the content of the Greek’s belief in the ordinary sense of the words just that there is lightning in his vicinity.}
out having justified beliefs concerning them)\textsuperscript{44}. Neither conceptual role with its verification conditions nor indication relations can account for the content of such beliefs or the meaning of sentences expressing them, so there is reason to think that Field’s later theory as a whole cannot account for them (which may be due to its holistic character)\textsuperscript{45}. However, his earlier more atomistic theory had at least some potential for dealing with such cases if further developed, as in it sentences speaking about unobservable entities could be constructed out of primitive terms referring to observable entities whose reference might be somehow explained causally, quite independently of the justification of the beliefs such sentences would express. Because of this Field’s earlier theory seems slightly better to me than his later one.

Furthermore, we can argue that physicalism or even moderate scientific realism cannot be coherently combined with strong kinds of deflationism, at least not with a pure disquotational theory, whatever its motivation. A physicalist (or any scientific realist) would surely have to say that many statements of current physical theory, or statements of an ideal physical theory, are true - this is generally taken to be what distinguishes physicalism from materialism pure and simple. However, modern physics uses lots of technical terms and outlandish notation, so that physicists (like other groups of scientists practising a highly developed science) essentially use a language of their own. Therefore people who are not (qualified to be) professional physicists cannot understand most of the sentences used in current

\textsuperscript{44}Also such beliefs can help us in predicting and explaining the behaviour of the person just as much as rational beliefs and their contents play an essential role in this; they can indeed not often explain the success of the person’s actions based on them, but they can explain and more importantly predict the failure of such activities. Predicting failure (and predicting exactly how a person is going to fail) is surely at least as important as predicting success, as failure is far more common; even if we cannot stop a person’s disastrous actions, preparing for their failure can help us to minimize the inevitable catastrophe.

\textsuperscript{45}One way for Field to respond to this argument would be to opt for eliminativism. He might say that instead of trying to construct our common notions of meaning and content out of conceptual roles and indication relations, we ought to just abandon those notions (as unacceptable for a physicalist) - which would involve abandoning the ordinary notion of belief itself, as the notion of content is surely essential to it - and replace them with conceptual roles and indication relations. Field does say that the line between reductionism and eliminativism is not sharp, so this would be a natural way to modify his theory. However, as deflationism has usually been advertised as a novel theory, it would become much less interesting for many philosophers if it were admitted that it just collapsed into old-fashioned eliminativism. Besides, Field has expressed reluctance toward accepting such eliminativist physicalism - e. g. he says in \textit{[Fie01b, page 30]} that we do unquestionably believe and desire - so there is a likely internal conflict in his theory as well as a conflict with common conceptions.
physical theory. Probably no one living today is capable of understanding an ideal physical theory, which would be likely to use far more technical terminology and complex notation. Therefore if anyone who is not a professional physicist or qualified to be one claims to accept physicalism, this must imply for him the claim that some sentences which he does not understand (or propositions he does not and often cannot grasp) are true. However, the pure disquotational theory demands that a person can only call such sentences as he himself understands true. Therefore if an adherent of a pure disquotational theory is not a professional physicist, as most deflationists surely are not, he cannot consistently say that such sentences are true and therefore he cannot be a physicalist (or even a moderate scientific realist). Since most philosophers are not (qualified to be) professional physicists and do not understand (and perhaps are not intelligent enough that they could ever understand) most of the sentences of a physical theory, they cannot consistently accept both a pure disquotational theory and physicalism (or even moderate scientific realism). Field could scarcely honestly say that he understands most of the sentences of physical theory, so if he were a pure disquotationalist he could not consistently be a physicalist. As I have shown before, it is not clear what kind of deflationist theory Field actually accepts in the end, so it is not clear if he could consistently be a physicalist, but there are surely at this point in the research good reasons to doubt (even if not definitely deny) whether any seriously deflationary conception of truth which differs significantly from more traditional inflationary correspondence theories is compatible with physicalism or even with moderate scientific realism.

4.3.2 Realism and the Theory of Truth

The relation of deflationism to the controversy between realism and antirealism in the theory of truth is as obscure as everything concerning deflationism. Indeed, the relations between the problem of truth and the problem of realism are subject to much dispute. Many philosophers, including many scientific realists, think that a correspondence theory is constitutive of realism, or at least of some kind of realism; e. g. William Alston defends in [Als97] alethic realism, realism concerning truth. If this is correct, and if deflationism is inconsistent with a correspondence theory, then deflationism would be inconsistent with realism. However, other philosophers such as
Michael Devitt in [Dev84] have argued that no theory of truth is constitutive of realism (though Devitt is still both a realist and a correspondence theorist). I have argued that Devitt is not consistent in this view. Sometimes deflationists (e.g. Horwich) agree with correspondence theorists like Devitt that the theory of truth is not relevant to the debate between realists and anti-realists, and this would make deflationism weaker than either realistic and anti-realistic theories. Horwich has defended this view strongly in [Hor90, §16-17].

However, there are reasons to think that even if deflationism does not directly and by itself imply anti-realism it is at least strongly motivated by anti-realistic considerations. As we have seen, Field presented the deflationist conception of truth as a way in which a verificationist can speak of truth. Verificationism, however, has traditionally been connected with anti-realism and there are reasons to think that it is inseparable from anti-realism. Field does indeed reject a verificationist definition of truth, but he yet seems to accept a verificationist theory of meaning, just as the logical positivists did. Also, of the two traditions which Field distinguishes, the one that he rejects, that of Frege and Russell, has traditionally been associated with realism.

Since deflationism is often motivated by verificationism, which is almost certainly incompatible with realism, a realist will be inclined to oppose deflationism. However, matters are complicated by the fact that deflationism seems to be sometimes motivated at least partly by other reasons than verificationism. We have seen that Field himself counted a deflationist, Ramsay, as part of the first tradition, the tradition that he did not say included verificationism (though there are surely verificationist elements in Ramsay’s Wittgenstein-influenced philosophy). Also when Field names verificationism as a crude paradigm of the second tradition, he leaves it unclear what less crude, subtler examples of the tradition might be. Also Field’s own sketch of the verificationist theory of meaning is in some ways weaker than the kind of verificationism favoured by logical positivists, so it might not suffer from all the same problems.

Horwich strongly opposes the idea that deflationism would imply anti-realism; however, there are reasons to think that Horwich may still himself be motivated by anti-realism. Horwich agrees with Field in rejecting a truth-conditional theory of meaning in [Hor90, §22, pages 71-74] and so apparently places himself in the second of the traditions distinguished by
Field: he says that meaning should be understood in terms of assertibility conditions rather than truth conditions, and assertibility conditions sound a lot like verification conditions. Even more, Horwich makes use in [Hor90, page 73] of Dummett’s manifestation requirement and acquisition argument. At the very least Horwich’s theory is incompatible with realism concerning understanding. Realism concerning understanding would assert that understanding is independent of (other) mental states, and this implies that it is possible that a person understands something even if others (or even he himself) cannot know that he does, and the manifestation requirement directly denies this.

However, we have seen that Dummett’s argument leads to a strong form of verificationism and Dummett argued that this kind of verificationism is incompatible with realism regarding pretty much anything. This certainly rouses a strong suspicion that Horwich’s overall semantic theory would have to be be inconsistent with realism regarding pretty much anything just like Dummett’s theory is, even if Horwich’s version of deflationism is not by itself inconsistent with realism. Moreover, Horwich’s verificationism is connected with deflationism; Horwich accepts in [Hor90, page 71] Dummett’s argument that deflationism is inconsistent with truth-conditional semantics. Therefore if the argument is correct (which, as we have seen, is highly controversial) deflationism lends at least some support for verificationist semantics (and hence for anti-realism), as it is the most prominent competitor for truth-conditional semantics, even if deflationism does not imply it outright.

However, Horwich refuses to admit that his theory would imply anti-realism. Is there then some difference in the kind of manifestation requirement Horwich accepts from the kind of manifestation requirement Dummett, the notorious anti-realist, accepts, that would explain and justify the different consequences the two philosophers draw from their seemingly identical premises? Or is there some error in Dummett’s argument that Horwich has seen? Unfortunately, it is far from clear that either would be the case.46

46Later writings of Horwich such as [Hor04] suggest there may indeed be at least one significant difference between Horwich’s sketch of a use theory of meaning and those of Wittgenstein, Dummett and Field, though it is not clear if those differences already existed in [Hor90]. Unlike Field, Horwich says (in [Hor04, page 357] that his attempt at a reductionistic theory of meaning is not motivated by physicalism. Horwich also says in [Hor04, page 358] that his version of the use theory of meaning leaves open accounts of meaning in psychological terms, and actually appeals to a visual experience in one of his examples [Hor04, page 351]. However, Dummett (and Wittgenstein) apparently
Horwich says in [Hor90, page 73] that he is not simply identifying the meaning of a sentence with its assertibility conditions. Unfortunately Horwich does not say much about exactly how meaning should be understood in terms of assertibility conditions if it is not to be identified with them, so unless he says more about his theory of meaning it remains unclear whether his theory can be combined with realism as he thinks it can.

If Horwich were correct that his theory does not imply anti-realism, then a defender of realism could accept Horwich’s kind of moderate deflationism and just try to make clear how his theory goes beyond it and argue that he is justified in going beyond it. In this case there would not be much difference between deflationism and weak correspondence theory for metaphysical purposes. However, as we have seen it is quite doubtful whether Horwich is correct, and therefore a realist is well advised to stick to a less deflationary correspondence theory.

4.3.3 Truth as a Property

The debate between deflationism and inflationism is also often connected with such questions as whether truth is a property; sometimes deflationism is characterised as denying that truth is a property. This kind of characterisation would bring the debate on the theory of truth into a close connection with the problem of universals, which will be the final theme of this dissertation. There are two extreme positions on this issue, between which I want to develop a compromise position. On the one hand some deflationists such as prosententialists deny that "true" is even a predicate, while on the other hand some inflationist correspondence theorists think that truth is a
sparse property. I want to show that neither of these extreme positions is attractive (though the second is arguably incoherent and the first not, so the first is not as bad as the second).

**Against Prosententialism**

Some deflationists such as prosententialists deny not only that truth is a property but even that "true" is a predicate. They claim (see e.g. [Gro92, pages 88-89]) that the truth term itself is a syncategorematic part of a prosentence, which is like an anaphor except that its syntactical category is a sentence.

The claim that atomic sentences in which the truth term occurs are prosentences does indeed sound plausible. However, this claim does not imply that the truth term would be syncategorematic. It is consistent to claim both that the truth term is a predicate and that atomic sentences whose predicate it is are prosentences. So what reason do prosententialists have to deny that the truth term is a predicate?

Prosententialists appeal to Russell’s treatment of definite descriptions, which according to Russell are not singular terms as they appear to be but quantifier phrases, as a precedent. However, few people today would say that Russell would have shown that definite descriptions and proper names in natural languages (which seem to be the main target of most deflationists, who seldom venture into constructing artificial languages - at least in detail - as Tarskians do) differ in their basic syntactic category (though some would say that they have different syntactic subcategorization features); both are clearly noun phrases. If there is a significant difference between them as Russell plausibly claimed, then this difference must be semantical, and manifest itself as syntactical only when natural languages are

---

47Grover has vacillated on this point. At one point she admitted that if having an extension sufficed for a property, then truth would be a property. However, as she tells the story in [Gro92, page 23], she then changed her mind because she became convinced that an extension does not suffice to establish that we have an interesting property. However, in this case we must remark that if such a property (an abundant property in Lewis’s terminology or a second class or third class property in Armstrong’s terms) is not interesting, then it is not interesting to deny its existence any more than to affirm it. Why then create a complicated semantical theory just in order to deny the existence of such an uninteresting property? In any case, also other philosophers such as extreme nominalists including Nelson Goodman and the early Quine in [GQ47] have expended great effort in trying to deny the existence of this kind of properties, which would seem to show that the claim that such properties exist is an interesting claim.
translated into a more perspicuous artificial language, where semantics and syntax are better correlated than in natural languages. However, there have been theories of natural language semantics that would remove even this difference; e. g. Montague’s theory takes both of them to be quantifier phrases (thereby radically altering our understanding of the semantic behaviour of singular terms such as proper names). Therefore the claim that ”truth” is not a syntactic predicate in natural languages is scarcely justifiable, but prosententialists like Grover do claim unlike Tarski that their theory of truth applies to natural languages like English. Also the fact that truth can be expressed by a prosentence in an artificial language does not imply that it could not be expressed by a predicate in another equally good artificial language, so prosententialism is not obviously incompatible with the kind of inflationism that claims ”true” to be a predicate in logical form. However, prosententialists at other passages merely deny that truth would be a property-ascribing predicate, and this is a slightly more plausible claim; it is a commonplace in ontological debates concerning the problem of universals that it can be held that at most some predicates, namely those that can be replaced by a variable bound by an objectual quantifier), ascribe properties to their subjects.

It is hard to see the motivation of this kind of radical deflationary theory. The picture it gives of the semantics of natural languages is highly unnatural even if coherent, so accepting it would require a very powerful argument. Though there are some disagreements among historians of philosophy with regard to the motivation of Russell’s theory of definite descriptions, Russell’s own claims indicate that it was at least partly motivated by the problem of non-existence. Russell wanted to eliminate apparent reference to non-existent and even impossible objects, in order to avoid a theory like Meinong’s theory. Whether or how far he succeeded is highly controversial; I argue in another section of this dissertation that there is also purported quantification over non-existent objects, which must also be somehow eliminated or accounted for, so Russell’s theory is not completely sufficient for his purpose, but is helpful for it. What then is the motivation of prosentential theories?

One prominent ontological argument against correspondence theories and therefore for deflationist theories such as prosententialist theories is that it is doubtful if correspondence theories are compatible with physical-
ism; however, it is not clear that prosentential theories would be compatible with physicalism either. It is not clear how an intralinguistic anaphoric relationship could be reduced into physical relations any more than how a relationship of reference between linguistic and non-linguistic items could be reduced to physical relations. Devitt for example claims in [Dev01b, page 604] that Brandom is totally upfront about abandoning naturalism and of course if this is correct it implies that Brandom also abandons the stronger theory of physicalism.

There is, however, a more serious problem with the prosententialist theory. Prosententialists claim that the role of the concept of truth is not to explain anything, but only to increase the expressive power of a language; a claim with which I am sympathetic, at least so far as it concerns the primary role of the concept of truth, since I am a mild deflationist. However, I will argue that they or at least most of them cannot provide an adequate account of this expressive role of the notion of truth, of how strong it is, which makes their theory too radically deflationist.

Prosententialists generally take the sentential quantification that they take truth to involve to be substitutional, rather than objectual quantification. This is because they often deliberately try to

---

48 Grover suggests in [Gro92, pages 113-115] that the prosentential theory can help a physicalist to get rid of mysterious facts concerning truth that do not seem to fit his world-view. However, she admits that this will leave him with the problem of facts about referring, facts about synonymy, and whatever other semantical facts there may be in the world, but claims that it is a start. However, Tarski’s theory as reformulated by the early Field could give the physicalist just as good a start, reducing truth to reference but leaving reference unaccounted for. Because of this Grover’s theory is in no way better than Tarski’s (as reformulated by the early Field) even from the point of view of a dogmatic physicalist, but requires a far more convoluted semantics than Tarski’s, so it is overall worse even for convinced physicalists. Brandom’s theory, which is far more radical than Grover’s, might help to get rid of reference but leaves anaphoric relationships unaccounted for, so it does not ultimately help to get rid of semantical facts and relationships either. Just like Tarski’s theory as reformulated by Field, these prosentential theories are just reducing semantical notions to other semantical notions, and doing nothing to reduce them to physical or other non-semantical ones. There is at present no guarantee that any theory of truth other than radical eliminativism would be consistent with physicalism, and many reasons to suspect that none is, though of course it is not certain that some less nihilist theory might in the end turn out to be consistent with physicalism either.

49 Mark Lance says in [Lan97, page 183, footnote 4] that one should not think that a substitutional interpretation of quantifiers is required of the anaphoric theory (of which the prosentential theory is a special case). There seem to be differences about where the boundaries of anaphoric theories lie. If an anaphoric theory allowed the use of objectual quantification, it might escape my objection. However, in this case it might no longer count as deflationary - at least, it would not be deflationary to the same degree as Grover’s theory is - and indeed Lance does claim that the anaphoric theory is compatible with
avoid ontological commitment to propositions, unlike such less radical de-
flationists as Horwich. However, this implies that a speaker of a language
can attribute truth only to sentences of the same language he speaks, since
only these sentences can serve as substitution instances of a sentential vari-
able bound by a substitutional sentential quantifier. He cannot attribute
truth to sentences in other languages, at least not those that could not be
translated into his language. However, it is surely possible to do this in
natural languages; for example, a speaker of English can say that something
a speaker said in German is true\textsuperscript{50}. Therefore Grover’s prosententialist the-
ory suffers from a similar defect as Field’s pure disquotational theory (but
which Horwich’s theory avoids, though only at the loss of clarity), where a
speaker could only attribute truth to sentences he himself understood. The
notion of truth Grover’s prosententialist theory yields is too weak to do jobs
a notion of truth is expected to do.

**Truth cannot be a sparse property**

Sometimes inflationism and deflationism are characterised so that inflation-
ism claims and deflationism denies only that truth is a genuine property
or that truth is a substantive property or a robust property etc. Unfortu-
nately it is often obscure what a genuine property or a substantive property
or a robust property is supposed to be, and also who is supposed to have
held that truth is a substantive property. A natural interpretation would
be that a genuine property is any property we are (or better ought to be)
tonologically committed to. However, often the expression seems to mean
something stronger. Often in such a case it means something we will have to
discuss at more length in the course of this dissertation, for example some-
thing like a sparse property in David Lewis’s sense or a first class property

\textsuperscript{50}This may indeed sometimes lead to paradoxes; however, there are cases where this
can occur without a risk of paradox (for instance, if the second language does not contain
any resources for talking about the truth of sentences in the first language or itself, but
still contains some non-semantic predicates which cannot be translated into the first lan-
guage). However, the mere fact that natural language can lead to paradoxes shows that
the notion of truth employed in it is radically stronger than the prosententialist notion;
mere sentential quantification could never lead to paradoxes. While the notion employed
by natural languages therefore has to be weakened in being regimented to avoid contra-
diction, this does not imply that it would have to be weakened so radically that it would
turn to anything like the prosententialist notion.
in Armstrong’s sense\textsuperscript{51}.

It seems to me that even if the theory of truth has significant metaphysical implications it need not follow that truth is a genuine property\textsuperscript{52} in any strong sense of that expression. These two views constitute then novel senses of the words “inflationism” and “deflationism”. It is blatantly false to claim that a correspondence theory of truth would require inflationism in this new sense, i.e. require truth to be a sparse property. In fact the claim that truth would be a sparse property is \textit{incompatible} with the basic idea at the basis of all correspondence theories of truth. As we have seen Wolfgang Künne has shown, all correspondence theories of truth take truth to be a relational property (if they take it to be a property at all); as I have also already shown, Bertrand Russell, who counts as a paradigmatic correspondence theorist for current philosophers, explicitly defined correspondence theory so that according to it truth was a relational property. However, David Lewis says in [Lew86, page 60] that sparse properties are intrinsic, and this can indeed be taken to be part of Lewis’s definition of a sparse property. That truth cannot be an intrinsic and hence not a sparse property

\textsuperscript{51}In [Asa11, page 152] Jamin Asay calls this kind of deflationism concerning truth metaphysical deflationism and explicitly identifies it with the denial of the view that truth is a sparse property. Asay is in favour of this metaphysical deflationism. However, Asay does not see that there is also a genuine metaphysical question concerning whether truth is an abundant property. He argues in [Asa11, page 177] that the property truth (by which he obviously means the abundant property truth) definitely exists, for there are truths (by which he means truthbearers). Asay seems to take that view as metaphysically innocuous; indeed, he says in [Asa11, page 146] following Shoemaker that there is a broad sense of the word “property” in which there is a property corresponding to any grammatical predicate. However, I have already argued that the postulation of abundant properties is not ontologically innocuous. There certainly is no sense of the word “property” in which there is a property corresponding to any grammatical predicate. However, I have already argued that the postulation of abundant properties is not ontologically innocuous. There certainly is no sense of the word “property” in which there is a property corresponding to any grammatical predicate, since the assumption that there is leads directly into a property-theoretical analogue of Russell’s Paradox. Consider the grammatical predicate “is an abundant property which is not instantiated by itself”; is it instantiated by itself or not? I certainly agree that there are truths, but it is important to see that this does not imply that an abundant property of truth would exist.

\textsuperscript{52}Devitt took in [Dev01b] the opposite position to the one Asay was to later take and opposed the adequacy of this kind of characterisation of deflationism. Devitt admits that deflationism has a metaphysical dimension and is a kind of anti-realism about truth and says that there is no reality to truth. However, Devitt thinks that this cannot be understood in terms of there being no property of truth, since in his view a nominalist or a selective realist like Armstrong who does not take truth to be a property need not yet be a deflationist. There are problems with Devitt’s attempt to separate realism concerning truth from realism concerning properties; for instance, Devitt does not notice that Armstrong distinguishes, as we will see, between first class and second class and third class properties, and truth (or at least the truth of sentences of a semantically open language) might well be a second class or third class property for Armstrong, even though it cannot be a first class property.
becomes still clearer when we look at Lewis’s way of clarifying the notion of intrinsic properties with the aid of the notion of possible worlds. According to Lewis [Lew86, page 62] an intrinsic property is one which can never differ between duplicates. However, any truth-bearer which is only contingently true or false (and does not concern truth-bearers) can obviously be true in a possible world and its duplicate can be false in another possible world. It is then certain that truth cannot be sparse property in Lewis’s sense.

The claim that truth is a sparse property fits better together with primitivist identity theories of truth (such as presented in [Moo99]), according to which truth is a primitive property of propositions or states of affairs such that a proposition or state of affairs with that property is a fact. Identity theories of truth, however, have generally been held to be incompatible with correspondence theories of truth (though they might also be counted as a limiting case of them, for identity is surely a limiting case of correspondence; everything corresponds with itself). The most that a correspondence theory (which is not an identity theory) might plausibly hold is that truth is an abundant or second class property.

There might be some interesting conception of ”genuine” property narrower than the concept of abundant properties other than Lewis’s concept of a sparse property that might apply to truth. However, it is not clear what it would be, and surely the burden of proof in on those who think that truth is a genuine property to define a concept of ”genuine” property that can do the job. In any case those ontologists who draw a distinction between ”genuine” properties and other properties generally think that relational properties are not ”genuine” properties. It is rather obvious that a relational property is nothing besides the relation on which it is based and it is easy to leap from this to the conclusion\footnote{Not everyone would leap to this conclusion. Aristotelians commonly tried to reduce relations to relational properties rather than the other way around. Some modern philosophers like Castañeda have defended this view; however, it is definitely a minority view, and its coherence is very controversial.} that it is not a genuine property.

It is not quite as clear that the correspondence relation or relations cannot be a sparse relation. However, presumably the same holds of sparse relations as of sparse properties namely that there are only just enough of them to characterize things completely and without redundancy. However, since the correspondence relation is according to correspondence theory to be defined with the aid of reference, then if reference is a sparse relation the
correspondence relation cannot be such, for in saying which terms are related by reference to which entities, we already characterize by implication all extensional semantical relationships completely. However, correspondence relations can be abundant relations, and I think that they are. In fact, I will later tell exactly what abundant relations these correspondence relations are, defining them with the aid of the relation of reference or denotation (leaving open whether this relation of reference or denotation is abundant or sparse).

However, neither correspondence theory as such nor Tarski’s specific theory implies even that truth would be an abundant property or that there would exist a correspondence relation as an abundant relation. In fact, Tarski’s original theory does not even require that we would have to take the truth-predicate as one of the expressions substitutable for a bound variable, much less for a variable bound by an objectual quantifier, so it does not even involve ontological commitment to truth (though very natural and minor extensions of it do). Therefore truth does not have to be a property at all, not even an abundant property, according to it. Nor does it seem to me that any weak definitions of correspondence theory would imply even this much; some such definitions may require that the correspondence relation used in the definition of truth is an abundant relation, but not that truth is a property; they can be taken to be just definitions of the truth predicate. Nor do all of them require even that the correspondence relation would be an abundant relation; the quantification over the correspondence relation many of them employ can be interpreted as non-objectual. Nevertheless, I think that the correspondence relations are abundant relations, which I will define, and therefore the version of correspondence theory I will develop at the end of this chapter in Section 4.4 will be slightly stronger, more inflationary, than Tarski’s theory, though it will still be far weaker than theories which would take truth to be a sparse property.

Indeed, there are reasons to think that even the supposition that truth is an abundant property might lead to semantical paradoxes (together with such plausible additional assumptions as the law of excluded middle). If such an abundant property of truth i. e. of being true as such, existed, and the law of excluded middle were true, then this property would have to be either instantiated by the Liar sentence or not, and both alternatives lead to a contradiction. However, I think that it is plausible to hold that
there are such abundant properties as being true in some semantically open fragment of English (in which phrase "some" has a broad scope); this claim cannot be adequately justified in this dissertation, but it becomes plausible once we have given arguments for the general claim that there are abundant properties.

4.3.4 The Theory of Truth and Explanation

The dispute between inflationists and deflationists is also often seen to revolve around the notion of explanation; deflationists often deny that the notion of truth would be explanatory and inflationists claim that it is. Those who hold correspondence theory to be inflationary by definition therefore often claim that correspondence theorists hold truth to be explanatory. The word "inflationist", the opposite of "deflationist", is often used as a synonym of "substantivalist", and for example Patterson says in [Pat03, page 422]:

As I use the terms a substantivalist account of truth is a theory on which truth is treated as a property that can in particular play certain explanatory role in serious theory. A correspondence theory of truth is a substantivalist theory that in particular treats this property as consisting in a relation between its bearer and something else, which relation in addition can be considered one of "correspondence"…

Substantivalism in this sense is one possible way of interpreting inflationism about truth, but a very strong way of interpreting even it. Not even versions of strong correspondence theory, such as fact-based correspondence theories, count as substantivalist theories of truth in this sense though they are surely inflationist in some sense. Fact-based correspondence theories such as Armstrong’s claim that truth is something to be explained rather than just analysed, but they need not claim that truth itself would explain anything; for them truth is an explanandum, not an explanans. It is far from obvious that this definition is equivalent with either of the ones I quoted previously, even though it shares some obvious similarities with it. It is not clear that an explanatory role is the same as a central role or a substantive role. A good feature in Patterson’s definition is that the notion of an explanatory role might be less vague than that of a central or substantive role, but it also suffers from unclarity, and in any case the expression ”serious theory”
introduces new vagueness of its own\textsuperscript{54}. It seems to me that any property (or other entity) whatever can occur in some explanations in some theories, and this might well mean that any property could play an explanatory role in some correct theories (and Patterson gives us no way to distinguish serious theories from frivolous theories) However, if fundamentalism is not correct, as I have argued in Section 2.1.4 of this work it might not be for all we know, then it might well be that no entity occurs essentially in any explanation in any correct theory. Any entity that occurs in an explanation might then be replaced by a more fundamental entity on which it supervenes. The notion of an explanatory role would then threaten to be trivialized, as either any entity would play an explanatory role or no entities would play such a role. Furthermore, there may be several notions of explanation. For these reasons Patterson’s definition does not make the matter sufficiently clear.

Field does say that full-fledged deflationism must deny that truth has an explanatory role, so an explanatory role seems to be a central role to him. Nevertheless, the concept of truth might play a central or substantive role in metaphysics without the notion of truth being explanatory; it might be that even if they do not explain anything, truths and truth conditions might still be something that would have to be explained and in this case they would have epistemological and metaphysical implications.

Many deflationists apparently accept a similar conception of correspondence theory as Patterson’s, and claim that their theory differs from correspondence theory since according to it truth has a purely logical role, not an explanatory one. Dorothy Grover apparently starts by supposing that an inflationary theory says that we need an analysis of the nature of truth, as she says in [Gro92, pages 4] that this is a common assumption of correspondence theories, pragmatic theories and coherence theories. She somehow infers from this that an inflationary theory implies that truth has an explanatory role, as she argues at length in [Gro92, pages 7-10] that the early Devitt is not correct in ascribing to truth an explanatory role and hence

\textsuperscript{54}At the end of his article in [Pat03, page 439] Patterson indicates that he is thinking primarily of a truth-conditional theory of meaning, so his notion of a substantivalist theory is in the end close to that of the later Field, which I have already discussed in section 4.3.1. Where the later Field opposes such a substantivalist theory, however, Patterson is in favour of it, and Field thought that Tarski’s theory could be developed to a substantivalist theory in this sense by being combined with a causal theory of denotation and so was at least compatible with such a theory while not amounting to such a theory by itself while Patterson thinks it was incompatible with it.
does not succeed in justifying a correspondence theory. It is not clear how she connects the two characterizations of an inflationary theory, for I cannot see why we should not need an analysis of the nature of an epiphenomenal property which has no explanatory role. Moreover, Dorothy Grover then goes on to remark (in [Gro92, page 189]) that logicians provide explanations. If logicians provide explanations, then a logical role can be an explanatory role, so this attempt to distinguish between correspondence theories and deflationary theories collapses into incoherence. One way to try to save this attempt would of course be to say that logicians provide a different kind of explanation than the one that correspondence theory claims the truth predicate to figure in: for instance, it might be suggested that the explanations in which the truth predicate figures are causal ones (as Field does seem to claim) while the explanations logicians provide are not causal. The problem with this suggestion is that it is not historically defensible; none of the paradigmatic correspondence theories claimed that the truth predicate would figure in causal explanations.

The problem with many arguments purporting to show that Tarski was not a correspondence theorist is that the same forms of argument could, if correct, be also used to prove that none of the paradigmatic correspondence theorists such as Aristotle or Russell or Reinach were correspondence theorists either. This is especially obvious when it comes to the arguments given by Patterson. Patterson at least addresses explicitly the question of what a correspondence theory of truth is, which many philosophers who treat the question whether Tarski was a correspondence theorist leave implicit. However, Patterson does not examine the history of how the term has been used in any detail and because of this his answer is not historically plausible.

While Patterson’s definition of a substantivalist theory is slightly vague, his definition of a correspondence theory is far worse. This latter definition is far too strong. If the notion of substantivalism is removed from his definition of a correspondence theory, we get what has been a standard account of a correspondence theory (as can be seen by comparing it to the definitions I have gone through). It seems to me that any theory of truth that treats the property of truth as consisting in a relation between its bearer and something.

---

55 It might be that Grover thinks that there could not be epiphenomenal properties, since she subscribes to some version of the Eleatic Principle according to which causal efficacy is a criterion of existence. I will argue in Section 5.3 that we have no reason to believe in such an Eleatic Principle.
else (which need not always be a truth-bearer) can be quite naturally called a correspondence theory of truth, even if it does not treat truth as a property that could play any explanatory role in any serious theory and even if it does not treat that relation in which truth consists as one that could play any explanatory role either.

This corresponds to the traditional definitions of correspondence theory that are found in the literature. As we have seen, neither Aristotle nor Russell stipulated that a correspondence theory would have to imply that truth can play any explanatory role. If we look at Crivelli’s thorough examination of Aristotle’s theory of truth, the notion of explanation does not once occur in it. Nor does Russell take truth to be an explanatory concept.56

Explanation was not the reason why these philosophers were interested in correspondence truth. Philosophers like Russell and Moore were chiefly interested in the correspondence theory of truth because it blocked the way to anti-realism and rank subjectivism that such competing theories as the coherence theory and the verificationist theory and the pragmatic theory.

56 Grover suggests in [Gro92, page 10] that Russell at least sometimes assumed that truth has an explanatory role, just like Field, since in his view (and Plato’s also) an account of knowledge called for an account of the difference between true and false belief. However, Russell was clearly thinking that the concept of truth was needed in the conceptual analysis of knowledge (as the overwhelming majority of philosophers have always thought and think even today, at least if they think that a conceptual analysis of truth is possible at all). However, this is entirely different from Field’s idea that truth has a causal explanatory role. After all, explanation and conceptual analysis are usually contrasted sharply, as we will later see in this dissertation Gonzalo Rodriguez-Pereyra does. As we have already seen, logical positivists such as Ayer thought that all analyses, the analysis of knowledge of course included, were tautologies. However, figuring as part of a tautology hardly makes truth a substantive concept! Even if we admit that all analyses are not tautologies and that conceptual analysis counts as explanation in a wide sense, as conceptual explanation, this is an entirely different kind of explanation from the kind of causal explanation Field had in mind. Because of this Grover gives no reason to think that Russell’s (or Plato’s!) theory of truth would be substantive in the same sense in which Field’s theory was. Indeed, it is even doubtful whether Russell could have thought that anything was to be causally explained in the literal sense, since Russell famously denied the existence of objective causal relationships in [Rus12a], though no doubt Russell could have given some substitute for causal explanation. Grover does seem to use the word “explanation” in a rather wide sense, since she says in [Gro92, §6.1, page 189] that logicians provides explanations; however, if she uses the word this widely, then it becomes problematic what reason she has to deny that truth would be explanatory in so weak a sense. However, Field clearly uses the word in a stronger, narrower sense. However, if we take Russell’s theory of truth to be a paradigmatic correspondence theory (as my reading of the literature indicates is usually done by those philosophers who bother to look at the history of philosophy to find out what a correspondence theory is supposed to be), this then implies rather conclusively that a correspondence theory of truth need not be a substantive theory of truth.
(which is closely allied to the verificationist one\textsuperscript{57}) offered. This does not imply that the correspondence theory would by itself ensure realism but only that it blocks some of the most important kinds of argument in favour of global anti-realism. We have already seen when discussing Devitt that realism implies the absoluteness of truth and correspondence theory helps to ensure this absoluteness. Versions of the coherence theory of truth (especially Neurath’s linguistic version in [Neu34]) and the pragmatic theory of truth are still around, and often even taken for granted, so opposition to them is far from irrelevant.

Patterson addresses in [Pat03, page 425] the notion of a weak correspondence theory\textsuperscript{58} formulated in [Wol] by Jan Woleński (who attributes the distinction between weak and strong correspondence to Kazimierz Twardowski) according to which a theory is a correspondence theory if according to it truth depends on extra-linguistic fact and argues that this is not a legitimate notion of correspondence theory. Patterson argues that even deflationary theories are correspondence theories in this sense and that therefore this is far too weak a notion to be of any interest if we wish to understand what is at issue between contemporary deflationists and their correspondence-theoretic opponents. It is of course quite plausible that this notion does not help us to understand that debate (though not all deflationists would agree that truth depends on extra-linguistic reality). However, the problem with this argument is that it assumes without justification that the notion of a correspondence theory should help us to understand what is at issue between these two camps. However, it is possible that the notion of correspondence belongs to an entirely different, older debate. Patterson is confusing, along

\textsuperscript{57}In [Jam48, page 160] William James said:

True ideas are those which we can assimilate, validate, corroborate and verify.

\textsuperscript{58}Patterson understands a strong correspondence theory so that it involves some generalized and noteworthy structural relationship between a sentence and the worldly state of affairs it concerns. This seems to restrict strong correspondence theories to fact-based correspondence theories, though Woleński’s original definition which Patterson quotes did not speak about states of affairs but only more generally of extralinguistic entities and so could count some object-based correspondence theories as a subspecies of strong correspondence theory. Patterson seems then to either misunderstand Woleński’s definition or alters it without giving any reason. In fact no plausible theory could hold that the kind of correspondence used in the definition of truth would be a purely structural relationship, though it might involve a structural component, since the notion of truth comprehends the notion of material truth, not merely logical or formal truth.
with many current philosophers, two very different debates. One of them is an old but yet still ongoing debate between correspondence theorists, coherence theorists, pragmatists - and let us remember that that as we have seen Tarski and other Polish logicians contrasted correspondence theory precisely with pragmatism - and other such camps. Weak correspondence does help us to understand what is at issue between these camps. The other is a newer, mostly unrelated debate between deflationists and inflationists. As I have already indicated when discussing Field’s development, the first of them is a relatively trivial debate about what truth is or how the truth predicate functions, but the other is a far more profound and difficult debate about the nature of more fundamental semantic notions like meaning and reference.

Patterson argues in [Pat03, page 432] for the same conclusion that the dependence of truth on extra-linguistic fact does not suffice to make a theory a correspondence theory also by considering a verificationist theory according to which sentences are associated with sets of experiences and are true if those experiences are had by someone. According to Patterson such a theory would not be a correspondence theory because all correspondence theories ought to endorse the principle that the things a sentence is about are among the things to which it corresponds; however, he says that a verificationist does not intend that all sentences are about experiences. I agree that a correspondence theory ought to endorse that principle (though endorsing it will not make it an explanatory theory)\(^{59}\). However, Patterson gives no evidence for his claim that a verificationist does not intend that all sentences are about experiences. As we have already seen in Section 3.3, historically most or all of the earliest verificationists (including Ayer and Schlick) were phenomenalists or neutral monists and did intend that all sentences would be about experiences, or if they did not consciously intend this, their overall theory in any case implied this. As we have seen, according to Ayer material and mental objects were logical constructions from sense-contents and it follows from this that sentences which apparently are about material or mental objects are really about sense-contents. However, sense-contents are

\(^{59}\)I agree that Tarski’s theory is slightly defective as it does not follow this principle; in the last section of this chapter of my dissertation (Section 4.4) I will suggest a small modification to Tarski’s theory that will remove this defect from it (without yet making it explanatory). However, I would say that a theory that did not endorse it would still count as a correspondence theory, only not a wholly satisfactory one.
what constituted experience according to Ayer, so even sentences ostensibly about material objects were according to Ayer all ultimately about experience. Later verificationists did indeed try to combine their theory with physicalism but it can be argued that this was not a consistent or well-considered combination. Verificationism can probably be combined with a correspondence theory, as. e. g. Moritz Schlick combined it (in Schlick 1929b and in Schlick 1934 - translated in Schlick 1959) and as Carnap tried to do. Such a verificationist correspondence theory would be a very unsatisfactory version of correspondence theory (but it would still be a better theory of truth than a coherence theory such as Otto Neurath’s in Neurath 1934, since it would be less subjective, as it would make truth independent of what we say or believe, even though not of what sensations we have or can have).

However, a correspondence theory of truth can also have a more constructive interest even if it does not treat truth as an explanatory concept, so it is far from being so uninteresting as Patterson paints it. Even if meaning cannot be explained in terms of truth-conditions, so that truth-conditions cannot serve as an explanans, yet truth-conditions can serve as an analysandum or an explanandum (even if only for a conceptual explanation, not a causal one) and therefore our knowledge of truth-conditions provided by a theory of truth can serve as evidence for a general theory of meaning, since a theory of meaning must at the very least get the truth-conditions of sentences right.

Patterson explicitly admits in Patterson 2003, 438] that his argument does imply that Aristotle was not a correspondence theorist (in any interesting sense) either; he does not note that the same argument probably applies to Russell and Moore and Wittgenstein etc. The earliest philosopher Patterson mentions whom he admits to have been a correspondence theorist is J. L. Austin (in an article reprinted in Austin 1961, though originally published in 1950). Even worse, however, Patterson does not give any reason to think that even Austin held truth to have any explanatory role, so even Austin’s theory need not count as a correspondence theory as Patterson has defined

\[\text{60}\] Patterson says that the idea of verificationism is to reconstruct being about cats from association with catty experiences, not to identify the two. This certainly does not hold of most verificationists. It holds to some extent of Field’s strongly revised modern version of verificationism (though Field uses in his reconstruction also such concepts as indication relations, which a verificationist could not accept). However, I have already argued that there are good reasons to doubt whether Field’s attempt to reconstruct content from verification conditions can work (even with the addition of indication relations).
it. In fact ordinary language philosophers like Austin generally held philosophy as a whole to be descriptive (though only of our linguistic practice rather than of the world as a whole or even of experience) rather than explanatory, so it is unlikely that Austin would have thought his theory of truth to be explanatory, though as far as I have found he does not explicitly state his position on this matter. Really, if Patterson were right about what correspondence theory is then Field would be the first correspondence theorist in the history of philosophy! This is of course an absurd conclusion; it is just as absurd as it would be to define dualism so strictly that Descartes was not a dualist (in any interesting sense) or redefining materialism so that Marx was not a materialist (in any interesting sense) or redefining theism so that Augustine was not a theist (in any interesting sense). Patterson is then just arbitrarily redefining the expression "correspondence theory of truth". Patterson’s argument is then in fact a reductio ad absurdum of the idea

we are told in a way independent of any assumptions about meaning under what condition in general a statement is true, namely, that the token state of affairs it is assigned by one set of linguistic conventions is of the type assigned to the sentence used in making the statement by another set of linguistic conventions.

However, none this implies that the property of truth would have any explanatory role - if this account makes anything have an explanatory role, it is conventions (and they would be explanatory only in a very weak sense of the word "explanatory"). Furthermore, in talking about conventions we are implicitly making quite strong assumptions about meaning, as it is conventions that give linguistic expressions their meaning. Saying that a state of affairs (or any entity) is assigned to a statement (or any linguistic expression) by linguistic conventions amounts to saying (as Austin himself stated plainly in a quotation Patterson himself gives on page 438 of his article) that the linguistic expression refers to (or denotes or means) it. Reference is, of course, what Field showed to be the basic concept in Tarski’s theory, so Austin made exactly as strong assumptions about meaning as Tarski did. Of course Tarski took for granted that reference or denotation was based on conventions. It is easy to reformulate Tarski’s theory in Austin’s terminology and this should show very clearly that Tarski makes no stronger assumptions about meaning than does Austin (but Tarski does not assume the existence of states of affairs as Austin does and therefore makes weaker assumptions overall). We have just to express the base clause of Tarski’s recursive definition in Austin’s words, as the rest follows from it. We can say that an atomic statement is true in general if and only if the sequence of entities which are assigned its arguments (i. e. its subject, object and indirect objects) by one set of linguistic conventions belong to the type of sequences assigned to its predicate by another set of linguistic conventions. You could try developing this into the germ of a reductive account of truth if you combined it with some detailed theory of conventions (such as David Lewis’s theory). However, no such theories of convention are (at least not yet) of a form acceptable to a physicalist (like Field), as all such accounts make use of mentalistic notions (though they might suffice the milder reductive ambitions of Horwich).
that Tarski was not a correspondence theorist and of the whole currently fashionable understanding of what correspondence theory is.

A more adequate view of the terrain is that there are both substantivalist and non-substantivalist correspondence theories. The notion of a substantivalist theory can be interpreted in different ways, so the distinction between substantivalist theories and non-substantivalist ones can also be understood in different ways, but on most ways of interpreting the distinction it is orthogonal to the distinction between correspondence theories and theories of truth which are not correspondence theories.

Patterson is at least partly right in one of his major claims, in denying that instances of the T-schema (which he calls T-biconditionals) state correspondences between the sentences they mention and something else. If we use the word "state" in a strong sense, then it is correct that the instances of the T-schema do not explicitly state any correspondences. If some correspondence theorists have stated that they would state correspondences, then they have been imprecise, though not far from the truth.

However, contrary to Patterson, Tarski did not assert that they state correspondences. In saying that the definition of truth is the logical product of the instances of the schema he is more likely intending to say that the theory of truth is implied, not explicitly stated, by such schemata. After all in talking about a logical product we usually talk about something that is derived by deductive inference from those sentences whose product it is.

Patterson, however, denies also that instances of the T-schema would even imply such correspondences. It is indeed not immediately obvious that they do, and it can be argued that Tarski’s own effort to derive correspondence from them does not wholly succeed. However, in order to derive correspondence relations from the instances of the t-schema we need just some kind of comprehension axiom schema of either higher order logic or set theory; I will show in the last section of this chapter in detail how this can be done.

---

62 Patterson argues that the T-biconditionals are not of the right form to state a relation at all, since they are just biconditionals. Tarski was of course quite aware of this, and this may be why he added the formal conditions of correctness to the requirements of an adequate theory of truth.

63 If the axiom needed is an axiom of set theory, then many would say that such an implication is not a purely logical one. However, I have already argued that is is far from obvious that set theory is not part of logic.
Patterson constantly confuses two questions: whether Tarski’s theory is a correspondence theory of truth and whether it is a satisfactory correspondence theory of truth. He also confuses the latter question with a third question; whether Tarski gives a good argument for a correspondence theory of truth. A theory of truth given by a philosopher could be a correct theory of truth even if the philosopher could not justify that it was. All the common definitions of a correspondence theory define not just one theory but a big family of theories; so there is not just one correspondence theory of truth, but there is an immense (probably infinite) number of possible correspondence theories of truth, of which presumably at most one is correct while all the others are false (with the caution that it may be hard to distinguish different theories from different formulations of the same theory, and therefore apparently different theories may turn out to be equivalent and so can both be correct). It is in fact highly unlikely that any philosopher will ever develop a fully satisfactory theory of truth, whether a correspondence theory or some other kind of theory; at most philosophers can hope to approach it. Therefore Patterson’s kind of fallacious argument would show that no theory of truth is a correspondence theory (or even that no theory of truth is a theory of truth at all). Patterson constantly gives arguments that could at most prove that Tarski’s theory is not a fully satisfactory correspondence theory (though most of them do not prove even that) and thinks that they prove that Tarski’s theory is not a correspondence theory at all. Perhaps the most egregious instance is in at the culmination of his argument in [Pat03, page 434]. In this place he starts from the correct claim that Tarski’s view amounts to a kind of extensionalism about truth, since the property of truth is identified as the property of being a member of a certain set. However, he then argues that we have been given no argument that truth itself is implicitly relational, unless we are given a nominalism about truth that few correspondence theories will accept. However, Patterson was not arguing just that Tarski did not give a satisfactory argument for correspondence theory but that Tarski’s theory was not even an unjustified correspondence theory. Therefore even if Tarski had not given any argument for the claim that truth is implicitly relational, the mere fact that he made the claim is a reason to think that he was a correspondence theorist, even if a bad one.

However, I do not think that Patterson has shown even that Tarski did not give a satisfactory argument for his correspondence theory. First
of all it is arguable whether Tarski’s extensionalism is really a nominalism, since most nominalists will not admit the existence of even sets which Tarski does\textsuperscript{64}. However, this is a purely verbal question; more importantly, it seems to me that the majority of correspondence theorists (and lots of philosophers who are not correspondence theorists, including deflationists like Quine) both at Tarski’s time and today do accept this kind of extensionalism, though this may be a small majority. At least Patterson does not provide any evidence that only few correspondence theorists would accept such an extensionalism. In any case since Tarski accepted such an extensionalism (and it was probably even a part of his conditions of formal correctness for truth) this provided him with a valid argument from the T-biconditionals and extensionality to a correspondence relation, and it is irrelevant that some correspondence theorists will not accept the premises of this argument.

An intensionalist correspondence theorist will think that Tarski’s theory is not yet fully satisfactory but not that it is a bad correspondence theory, much less that it would not be correspondence theory at all. Besides, even many intensionalists about properties will usually claim that every property has a set as its extension (as e. g. Carnap did as one of the first). There are many exceptions, like the early Goodman, but I doubt that they are a majority. Therefore most intensionalists will agree that a complete theory of truth, even if it treats truth as an intensional property (or at least as an intensional predicate), will have to tell which is the extension of that property (or predicate), and this will still require treating truth as a relational property, if Tarski’s argument is valid (and Patterson gives no reason for supposing that it would not be). Therefore a very great number of philosophers, whether extensionalists or intensionalists, should accept Tarski’s argument for the relational character of truth as valid.

It may be a relatively uninteresting verbal question what we should mean with the expression “correspondence theory of truth”; while I have argued that the way many philosophers like Patterson use the term today is highly misleading and therefore inappropriate, it is not of course wrong in any absolute way, since language is conventional and any term can be used any way

\textsuperscript{64}There is evidence that the late Tarski denied the existence of sets in conversation. However, this does not show that he would not have believed in them when formulating his theory of truth. In any case, if he had denied the existence of sets as a metaphysical thesis while yet making use of sets, this would only show that he was so far inconsistent.
one pleases. However, it is interesting that the way the terminology is generally used and the debate conducted hides out of sight one very interesting option in the theory of truth, the view that truth is a complex relational property (or at least the truth term is a one-place relational predicate) but truth has no explanatory, at least no causally explanatory role (i.e., is epiphenomenal). I think that it is most appropriate to call such a theory a non-substantivalist correspondence theory of truth, but since language is conventional, it could of course also be called something else. What I am arguing is that this kind of theory should be seriously considered; it seems plausible to me, even though I am not willing to yet fully commit myself to it.

This is basically because of doubts which I have already expressed about what counts as an explanation and explanatory role. (There are even more difficult questions about what counts as a causal explanation and as a causally explanatory role, but I do not have space here to go into the causal aspect of the question.) Though the notion of explanation is heavily used in modern philosophy, detailed theories of what explanation is are rare yet exhibit remarkable variety given their small number. The notion of truth can of course be used in some low-level or proximate explanations; any notion can be used in some explanations. If claiming that truth is an explanatory notion is to help to erect a dividing line between (some kinds of) inflationism and deflationism, it surely must imply something more, such as that the property of truth occurs essentially in explanations. Whether it does so occur is a different and harder question. Putnam and the early Field did give a plausible (and very simple) argument for the claim that the notion of truth has an explanatory role; the fact that a person has true beliefs surely often helps to explain his success in his projects. However, both later came to reconsider the correctness of this argument.

The late Field argues against this earlier theory of his own in his last postscript to [Fie94] in [Fie01b, pages 153-156]. His argument seems to boil to the inference that truth does not occur essentially in explanations, and hence is not explanatory (though Field does seem to admit that truth can occur in a second class explanation, so presumably it can be still said to play a second class explanatory role). Field argues that explanations of a person’s ability are expressible (in principle) without talking of the truth conditions of a person’s representations, and infers from this that reference to
truth conditions plays only a heuristic role. However, as I have argued, if ontological fundamentalism is false, then it might be that no property or particular occurs essentially in any explanations. In that case any explanation could be replaced with a fuller explanation, so all explanations would be second class explanations. Therefore this does not show that truth is any less explanatory than any other entity. Furthermore even if fundamentalism turned out to be true, any explanation making use of familiar middle-sized is in principle expressible without talking of these entities and talking only of their microscopic constituents. Therefore an argument of the very same form as Field’s argument could be used to show that no macroscopic entities have an explanatory role but of entities known to us only such as quarks and electrons and their properties could do so, and macroscopic entities would play only a heuristic role and hence that any macroscopic properties and particulars are epiphenomenal (or, as mereological nihilists claim, do not exist at all). However, it seems absurd to claim that no macroscopic entities play an explanatory role (or that any theory concerning them would be a frivolous theory) and still more absurd that there would be no macroscopic entities, though there have been philosophers, extreme scientific realists, who have defended this conclusion.

Therefore if we are willing to admit that any familiar middle-sized entity whatsoever plays an explanatory role, we should admit that truth also plays an explanatory role and that our account of truth should be substantivalist (at least on some ways of interpreting Patterson’s imprecise definition). Nevertheless, that truth is explanatory need not be taken to be part of the definition of correspondence theory, since a correspondence theorist surely could consistently believe that all macroscopic entities are epiphenomenal. Furthermore, that truth is in some way explanatory does not imply that the explanatory role of truth would be the only or primary reason why we need the concept of truth; I agree with the deflationists so far that I think the primary reason why we use a concept of truth is for expressive purposes.

4.3.5 Conservativeness and Deflationism

Jeffrey Ketland suggests in [Ket99, page 79] that non-substantiality is conservativeness from which suggestion it of course follows that substantiality is non-conservativeness. Ketland further distinguishes in [Ket99, page 71-72] two kinds of conservativeness, semantical conservativeness and deductive
conservativeness and claims that the non-substantiality of deflationist theories of truths is deductive conservativeness.

Ketland’s suggestion is interesting but there are some problems with it. Intuitively substantiality is a general notion which should be applicable to other theories than the theory of truth. However, there are serious problems with extending Ketland’s identification of substantiality with non-conservativeness to other subjects. It seems intuitively that "substantiality" is a monadic predicate. However, "conservativeness" is a dyadic one; as Ketland says, a metatheory is conservative over theories in the object language or in other words a theory is a conservative extension of another theory. Because of this conservativeness cannot serve as a general criterion of substantiality, even if it might do so in the special case of theories of truth. However, the claim that it does so in the case of truth would require a deeper motivation than Ketland gives.

A deeper worry lies in Ketland’s appeal to deductive conservativeness. After all, the concept of truth is a semantical one, so surely the distinction between two theories of truth should be semantical, not syntactic, but the concept of deductive conservativeness is syntactic. Trying to defend Ketland’s view, Shapiro claims in [Sha98, page 499] that in order to capture generalizations concerning truth the deflationist must appeal to a consequence relation that is higher order while the Tarskian can use a first-order consequence relation. However, Tarski’s original theory of truth was a higher-order theory, so Tarski had in any case to appeal to a higher-order consequence relation! It must of course be admitted that later modifications of Tarski’s theory have usually (though by no means always) been first-order theories. However, surely Tarski’s original theory should also count as a correspondence theory. There is the danger in Shapiro’s argument that it will imply that Tarski’s original theory was non-substantial i.e. deflationary and only later modifications of it count as substantial, which was surely not Shapiro’s intended conclusion. Furthermore, the reason philosophers are often opposed to a higher-order consequence relation is that it is undecidable. However, if as I argue correspondence theory is motivated by realism, then a correspondence theorist should in any case be hospitable to undecidable notions. After all, a realist should have no problem with undecidable relations any more than with any other things which are not knowable completely or easily; it is only anti-realists and extreme verificationists of Dummett’s type
who are hostile to undecidable notions.

Ketland tries to show that various deflationary theories of truth are deductively conservative over theories the truth of whose sentences they concern while Tarski’s theory is not. Ketland’s proof may indeed show that many deflationalist theories (even if not necessarily all) are because of their conservativeness weaker than Tarski’s theory and are therefore at least in one respect more deflationary than it, i.e. that Tarski’s theory is in one respect more inflationary than it. This fits well with my view that Tarski’s theory is only mildly inflationary while theories called deflationary by their proponents are strongly deflationary. However, I do not think that Ketland’s proof would show that we could establish a binary distinction between deflationary and inflationary theories with the aid of the concept of conservativeness.

Stewart Shapiro claims in [Sha98, pages 503-507] that non-conservativeness also implies being explanatory (a claim that goes beyond what Ketland says in [Ket99]). Shapiro admits in [Sha98, page 507], however, that there is no consensus on the notion of explanation, and that the notion he requires is one on which there is explanation in mathematics, while other accounts require a causal relation between explanans and explanandum. He insists, however, that on an intuitive level a Tarskian theory of truth is explanatory. I would suggest, however, that have to do here not with competing accounts of one notion, but with several basically different notions of explanation. In one very weak sense of the word ”explanation” mathematics does provide explanations, but this is a very different sense of the word than the commonly used one in which theories in natural science provide explanations (often but not always causal ones). Therefore while Shapiro’s argument shows that Tarski’s theory is inflationary to some degree, it would still allow it to be less inflationary than Field’s original theory. My argument at the end of Section 4.3.4 of this work shows that truth does play a role at least in second class explanations of behaviour, so if correct it shows the notion of truth to be more inflationary than Shapiro’s, but this is still less inflationary than the kind of theories philosophers like Patterson and Sher dream about.
4.4 A New Formulation of Correspondence Theory: Non-dyadic Correspondence Relations

Künne’s characterization of object-based correspondence theories of truth has one serious limitation. Künne supposes that the implied relation has to be a two-place relation, whether a relation between sentences and states of affairs as in fact-based correspondence theories or a relation between the referent of the subject and the predicate as in object-based correspondence theories. However, Künne’s own definition (or Russell’s formulation) does not imply this. The notion of correspondence is often outside the theory of truth (e. g. when discussing intertheory relations between theoretical terms) used in logic and philosophy for relations which are not dyadic, so there is no reason why it would have to stand for dyadic relations in theory of truth either. It seems to me that in order to satisfy this definition and the realistic intuitions behind correspondence theory that it expresses, all that is necessary is that the implied relation is a polyadic relation and the first member of the implied relation must be a truth-bearer but the remaining members of the relation, no matter how many of them there are, need not be truth-bearers.

Even more, there need not be a single correspondence relation, but there can be different correspondence relations for different sentences. A correspondence theory says that a truthbearer is true if and only if it corresponds to reality, i. e. if and only if for every true truthbearer, there is a correspondence relation (an abundant, not a sparse relation) between it and a reality. There is no reason why this relation would have to be the same in the case of all sentences or all judgements or all propositions. There is, however, a reason why this relation cannot be the same; supposing a single correspondence relation leads swiftly to paradoxes. Of course, these relations must have something in common since they all are relations on which truth depends and presumably the truth of different truthbearers is somehow similar.\textsuperscript{65}

In fact Russell in his theory considered the possibility that we might

\textsuperscript{65}It might be possible to make do with one single correspondence relation if it were taken to be an anadic relation, as in the anadic logic developed in [Gra76]; however, I will not investigate this possibility at further length but will develop my new version of correspondence theory in a predicate logic closer to standard predicate logic.
avoid having to postulate objective propositions\textsuperscript{66} as correlates of belief (as he himself had done earlier) by supposing that belief was a more than two-place relation. According to him, when Othello believes that Desdemona loves Cassio, then there is a relation of four terms whose members are Othello, Desdemona, Cassio and loving. However, Russell still thought that in the case of a true belief the objects of the belief formed a complex that was a fact, e. g. in the case of the previous example, Desdemona’s love for Cassio. However, Russell’s train of thought can be followed further by taking also the relation implied by the relational property of truth to be a more than two-place relation, similar to the relation of belief in Russell’s theory (though it obviously cannot be the same relation, for innumerable reasons, perhaps the most important of which is that there are false beliefs and in order to be false they cannot correspond to anything in the same way in which true beliefs do), and so it can lead us to consider eliminating facts just as easily as propositions. Therefore when the sentence or belief ”Desdemona loves Cassio” is true, then a four-place relation holds between the sentence or belief, Desdemona, Cassio and loving, the relation that holds between four objects when the first of them is a truthbearer consisting of three terms denoting the three remaining objects and the fourth of the objects is a relation which holds between the second and third object. In symbols, if we take $C(x_1,x_2,\ldots,x_n)$ to mean that $x_1$ is the concatenation of $x_2,\ldots,x_n$ and $D(x,y)$ to mean that $x$ denotes $y$ (i. e. that $y$ is assigned to $x$ by some linguistic conventions, to use Austin’s and Pattersons’ terminology), then the relation is $(\lambda x)(\lambda y)(\lambda z)(\lambda R)((\exists u)(\exists v)(\exists w)(C(x,u,v,w) \land D(u,y) \land D(v,z) \land D(w,R) \land R(y,z)))$.

This would be easy to generalize to all sentences if it were not for the paradoxes; in a naive theory of truth where we make no careful distinction between a language and its metalanguage we would just say that if $\phi$ is a sentence with $n$ constant terms (that do not occur in an intensional context) $t_1,\ldots,t_n$ and $C$ is such a relation that for all terms $x, x_1,\ldots,x_n$ $C(x,x_1,\ldots,x_n)$ iff $x = \phi(x_1/t_1,\ldots,x_n/t_n)$ where $\phi(x_1/t_1,\ldots,x_n/t_n)$ is the

\textsuperscript{66}Russell still used the word ”proposition” in this writing, but this is very misleading since he only means declarative sentences with the word, reverting to pretty much the medieval terminology after he had been one of the philosophers who trained analytical philosophers of his time to use the word in a very different way. What he denies are objective falsehoods such as ”that Desdemona loves Cassio” in the case in which Desdemona does not love Cassio. Clearly these objective falsehoods are what are usually called false propositions.
formula which we get when we replace all the constant terms (which occur in an extensional context) with variables then the implied relation is 
\((\lambda x)(\lambda y_1)\cdots(\lambda y_n)(x = \phi \land D(t_1, y_1) \land \cdots \land D(t_n, y_n) \land \phi(y_1, \ldots, y_n)).\) In order to take account of the semantical paradoxes we must make our theory more complex; if \(\phi\) is a sentence with \(n\) constant terms (occurring in an extensional context within it) \(t_1, \ldots, t_n\) and \(M(\beta_1, \ldots, \beta_n)\) is the translation into the metalanguage of the formula \(\phi(x_1/t_1, \ldots, x_n/t_n)\) which we get when we replace all the constant terms with variables the implied relation is 
\((\lambda \alpha)(\lambda \beta_1)\cdots(\lambda \beta_n)(\alpha = \phi \land D(\delta_1, \beta_1) \land \cdots \land D(\delta_n, \beta_n) \land M(\beta_1, \ldots, \beta_n)).\)

There is also the far more serious problem that this account might lead to set-theoretical (or rather property-theoretical) paradoxes; if \(\phi\) is \(Q(Q)\) or \(Q((\lambda x)(Q(x)))\) or \(t_i \in t_i\), then the correspondence relation so defined cannot exist on pain of contradiction, at least in a two-valued logic. However, we can easily restrict these relations further in some way so that paradoxes cannot arise. There are many ways to do this; as a very simple but rather extreme solution, we can define a separate correspondence relation for every sentence (where the earlier theory only defined different correspondence relations for sentences of different logical forms). The correspondence relation for many sentences, especially unquantified positive sentences, will then in such a theory be a relation that is instantiated by a single sequence if the sentence in question is true and by no sequence (i.e. will have the empty set as its extension in a normal set theory) if the sentence is false. If \(\phi\) is a sentence with \(n\) constant terms (which occur in it in an extensional context) \(t_1, \ldots, t_n\) and \(M(\beta_1, \ldots, \beta_n)\) is the translation into the metalanguage of the formula \(\phi(x_1/t_1, \ldots, x_n/t_n)\) which we get when we replace all the constant terms with variables then the implied relation is 
\((\lambda \alpha)(\lambda \beta_1)\cdots(\lambda \beta_n)(\alpha = \phi \land D(\delta_1, \beta_1) \land \cdots \land D(\delta_n, \beta_n) \land M(\beta_1, \ldots, \beta_n)).\)

If it can be understood as an extensional relation then it is the same as the class \(\{\langle \alpha, \beta_1, \ldots, \beta_n \rangle : \alpha = \phi \land D(\delta_1, \beta_1) \land \cdots \land D(\delta_n, \beta_n) \land M(\beta_1, \ldots, \beta_n)\}\); even if understood as an intensional relation, however, it has this class as its extension. Though this kind of theory will give every sentence its own correspondence relation (which may or may not hold), it is yet a general theory, since it gives a general rule by which all of these correspondence relations can be constructed for every sentence.

In fact this formulation makes use of unnecessarily many primitives and
indeed sounds yet too deflationary, since while it makes use of the notion of denotation it also uses the interlinguistic notion of translation as a primitive. The notion of denotation must according to realist theories be prior to the notion of translation, as denotation is a relation between linguistic expressions and the rest of the world while translation is a relation between linguistic expressions, and a realistic theory takes relations between language and the world to be prior to any semantical relations between linguistic expressions. Therefore it not yet a definition of truth but only a material condition of adequacy for a definition of truth like the T-schema (though closer to a true definition) and it must be replaced with a recursive definition of correspondence relations analogous to Tarski’s recursive definition of truth itself.

I will begin with defining correspondence relations for an extensional language, because I can then follow the Tarskian example rather closely. I will suppose for the sake of simplicity that all primitive singular constants in the language that are of a kind that could denote actually denote something. This will have to supplemented later with a definition for intensional languages, in which the non-linguistic members of the correspondence relations are at least in some cases intensions or characters or other semantic values rather than extensions of the constants of the language. We might eventually have to assume that some sentences such as those containing non-denoting singular constants are neither true nor false, and in this case they would not have any correspondence relations associated with them.

I will then give the following recursive definitions of correspondence relations and of truth itself. I will use \( \text{Corr}(\phi, g) \) as the symbolization for the correspondence relation for sentence \( \phi \) with respect to assignment \( g \).

**Definition 1** A sentence \( \phi \) with is true if and only if there are such entities \( \alpha_1, \ldots, \alpha_n \) (for some number \( n \)) that for all assignments \( g \) \( \text{Corr}(\phi, g)(\alpha_1, \ldots, \alpha_n) \).

The definition of correspondence relations must naturally begin from atomic sentences.

**Definition 2** If \( \phi \) is identical with a sentence \( R(t_1, \ldots, t_n) \) and there is such an entity \( \alpha \) that \( \|R\| = \alpha \), then \( \text{Corr}(\phi, g) \), the correspondence relation relative to assignment \( g \) for it is the \( n \)-place relation \( (\lambda \alpha)(\lambda \beta_1) \cdots (\lambda \beta_n)(\lambda \gamma)(\alpha = \phi \land \|t_1\|_g = \beta_1 \land \cdots \land \|t_n\|_g = \beta_n \land \gamma = \|R\|_g \land (\beta_1, \ldots, \beta_n) \in \gamma) \).
which has as its extension the class \( \{ (\alpha, \beta_1, \ldots, \beta_n, \gamma) : \alpha = \phi \land \| t_1 \|_g = \beta_1 \land \cdots \land \| t_n \|_g = \beta_n \land \gamma = \| R \|_g \land \langle \beta_1, \ldots, \beta_n \rangle \in \gamma \} \).

Accounting for the truth of negative sentences has usually been very difficult for correspondence theories, especially fact-based ones and most especially truth-maker theories. Especially negations of existentially quantified sentences - of the form \( \neg (\exists x) \phi \) - are problematic. One of the most intense debates between truth-maker theorists has been whether negative facts are needed as truth-makers for negative sentences. Even an object-based correspondence theory, though free of this problem, would run into troubles with the case of negative sentences if it was based on the assumption that the terms which correspond were singular terms and the members of the correspondence relation on the mundane side would have to be particulars, since in this case the correspondence relation would in the case of negations of existentially quantified sentences collapse to a property of the truth-bearer. However, even a negation of an existentially quantified sentence always contains some predicate terms to which abundant properties will correspond; e.g. \( \neg (\exists x) P(x) \) contains \( P \), \( \neg (\exists x)(P(x) \land Q(x)) \) contains \( P \) and \( Q \) etc. There might also be a problem if the correspondence relations (or the relations that might be their members) were taken to be sparse properties, for the correspondence relations for negative sentences would be negative relations; the correspondence relation for a negative sentence would be a relation which would hold between the sentence and the entities which the terms occurring in the sentence denote if they did not satisfy a condition, namely the condition expressed by the sentence whose negation the negative sentence in question is. However, negations of sparse relations are according to most theories of sparse properties not sparse relations. However, if abundant relations are taken as correspondence relations there is no problem since negative relations are quite legitimate abundant relations.

**Definition 3** If \( \phi \) is identical with a sentence \( \neg p \) and the correspondence relation for \( p \) relative to assignment \( g \) is an \((n + 1)\)-place relation \( R \) then the correspondence relation for \( \phi \) relative to \( g \) is the \( n + 1 \)-place relation \( (\lambda \alpha)(\lambda \beta_1) \cdots (\lambda \beta_n)(\alpha = \phi \land \neg (R(\phi, \beta_1, \ldots, \beta_n))) \) which has as its extension the class \( \{ (\alpha, \beta_1, \ldots, \beta_n) : \alpha = \phi \land \neg (\phi, \beta_1, \ldots, \beta_n) \in R \} \).

**Definition 4** If \( \phi \) is identical with a sentence \( p \land q \) and the correspondence relation for \( p \) relative to assignment \( g \) is the \((n + 1)\)-place relation \( R_1 \) and
the correspondence relation for $q$ relative to assignment $g$ is the $(m+1)$-place relation $R_2$ then the correspondence relation for $ϕ$ relative to $g$ is the $(n + m + 1)$-place relation $(λα)(λβ_1)⋯(λβ_n)(λγ_1)⋯(λγ_m)(α = ϕ ∧ R_1(p, β_1, ⋯, β_n) ∧ R_2(q, γ_1, ⋯, γ_m))$ which has as its extension the class $\{⟨α, β_1, ⋯, β_n, γ_1, ⋯, γ_m⟩ : α = ϕ ∧ R_1(p, β_1, ⋯, β_n) ∧ R_2(q, γ_1, ⋯, γ_m)\}$.

**Definition 5** If $ϕ$ is identical with a sentence $p ∨ q$ and the correspondence relation for $p$ relative to assignment $g$ is the $(n+1)$-place relation $R_1$ and the correspondence relation for $q$ relative to $g$ is the $(m + 1)$-place relation $R_2$ then the correspondence relation for $ϕ$ relative to $g$ is the $(n + m + 1)$-place relation $(λα)(λβ_1)⋯(λβ_n)(λγ_1)⋯(λγ_m)(α = ϕ ∧ (R_1(p, β_1, ⋯, β_n) ∨ R_1(q, γ_1, ⋯, γ_m)))$ which has as its extension the class $\{⟨α, β_1, ⋯, β_n, γ_1, ⋯, γ_m⟩ : α = ϕ ∧ (R_1(p, β_1, ⋯, β_n) ∨ R_1(q, γ_1, ⋯, γ_m))\}$.

Some might see it as a defect of this definition that it allows arbitrary entities to appear as members of the correspondence relations. We could avert this by instead using the following definition: If $ϕ$ is identical with a sentence $p ∨ q$ and the correspondence relation for $p$ relative to assignment $g$ is the $(n+1)$-place relation $R_1$ and the correspondence relation for $q$ relative to $g$ is the $(m + 1)$-place relation and $r = max(n + 1, m + 1)$, then $R_2$ then the correspondence relation for $ϕ$ relative to $g$ is the $r$-place relation $(λα)(λβ_1)⋯(λβ_n)(α = ϕ ∧ (R_1(p, β_1, ⋯, β_n) ∨ R_1(q, β_1, ⋯, β_m)))$.

**Definition 6** If $ϕ$ is identical with a sentence $(∃x)ψ$ then the correspondence relation for $ϕ$ relative to assignment $g$ is $(λα)(λβ_1)⋯λβ_n)(α = ϕ ∧ (∃x)(Corr(ψ, g(γ/x))(ϕ, β_1, ⋯, β_n)))$ which has as its extension the class $\{⟨α, β_1, ⋯, β_n⟩ : α = ϕ ∧ (∃γ)(⟨ϕ, β_1, ⋯, β_n⟩ ∈ Corr(ψ, g(γ/x)))\}$.

This is obviously not yet a wholly rigorous development of such an object-based correspondence theory, though it already comes close. It would be rather complicated to develop this kind of account fully rigorously, but it could be done. It may be not obvious that it is worth doing, since it would just mean developing a more convoluted notational variant of Tarski’s theory. The theory I have presented is of course equivalent with Tarski’s theory given some very weak set-theoretical assumptions; however, this does not imply that it does not add anything to Tarski’s theory (after all, it has been argued - though the validity of the arguments is controversial - that some deflationary theories are also equivalent with Tarski’s, yet they have been
held to have different philosophical implications). Such a notational vari-
ant can be philosophically illuminating, since it would show explicitly and
in detail how Tarski’s theory is a version of the correspondence theory in
the sense of Russell’s definition even though it does not postulate facts or
truthmakers in the standard sense.
Chapter 5

A Theory of Ontological Commitment

5.1 Introduction

So far I have argued that neither verification conditions nor truthmakers are central to the ontological analysis of scientific and other theories. While the arguments by which I have tried to show this are not fully conclusive, I think that they have at least motivated a tentative search for some alternative to both. This has been the negative part of my thesis, and now I can proceed to a more positive part it, which many readers may find more interesting. I will argue that ontological commitment is the central metaontological concept, and try to develop a fully rigorous theory of it. However, there yet remain some negative arguments, as I must try to overthrow views of ontological commitment I find to be false.

This chapter deals with the metaontological question of what ontological commitment is. The question of ontological commitment is what we assume to exist, i.e. are ontologically committed to, in our discourse. Obviously if we cannot answer this question, we may then both implicitly suppose that there are some entities and implicitly deny this without being able to notice the inconsistency. In such a case we are unable to find out what evidence our theories outside ontology provide for ontological theories. Therefore this is one of the most important metaontological questions, one which must be answered before any substantive ontological questions can be addressed in any very rigorous way. The fact that ontologists today often disagree greatly
over what ontological commitment is shows that at least many of them do not really know clearly what they are doing. Ontologists rely on their intuitions about what ontological commitment is, and the intuitions of different ontologists do not agree even to the very limited extent that the intuitions of philosophers in other fields of philosophy agree. This may be one of the reasons which explain the relative lack of progress in ontology. Because of this it is of the utmost importance for the future progress of ontology to develop a systematic and rigorous theory of ontological commitment, a theory that does not depend on any direct ontological intuitions. Naturally such a theory will have to depend on intuitions at some point, to gain some sort of starting point, but I will argue that these need not be directly ontological intuitions, but can be logical and mathematical and linguistic, specifically semantical intuitions. Since logic and mathematics and linguistics have (in the opinion of most philosophers) made far more progress than ontology, this should be a key to progress in ontology also, if it is possible.

I attempt to construct such a systematic theory in this chapter of this dissertation by retrieving forgotten ideas of Alonzo Church and combining them with the basic ideas of Tarski’s theory of truth. The resulting theory of ontological commitment will be just as rigorous as Tarski’s theory of truth and its modern day descendants, such as the model-theoretic conception of truth. As a price of the rigorousness it will naturally be rather trivial, just as Tarski’s theory has been accused of being. Of course, in spite of its rigour it will still not be fully certain, since even Tarski’s theory is not that; as I will show, nearly all of the objections that have been commonly brought against Tarski’s theory would also, if there were anything to them, apply to the theory presented in this dissertation. Therefore those who reject Tarski’s theory of truth (as not only trivial, but as wrong) will naturally, if they are consistent, also reject the metaontological theory of this dissertation and therefore also its substantive ontological consequences. However, it will be rather hard for those who wholly accept Tarski’s theory of truth to reject the theory presented in this dissertation.
5.2 Historical Background: Earlier Candidates of the Criterion of Ontological Commitment

5.2.1 Logical Positivism and Reality and Existence

The most famous answer to the question of the criterion of ontological commitment was that of Quine, who famously proposed in [Qui53d, page 12] that we are committed to what we quantify over with bound variables.

In order to see the importance of Quine’s criterion properly we must contrast it with earlier answers given by analytical philosophers during the dominance of logical positivism to the problem. For instance, the logical positivist Schlick tried to define actuality, reality or existence (all three of which concepts he apparently presupposed to be the same)\footnote{It is plausible that these words “actual”, “real” and “existing” are synonymous on at least one reading of them. However, it is not clear that this is the case with all readings of them, and it may be that in some sense of the words actuality and reality are more than just being a value of the bound variable. The German word Schlick and Carnap used, which is translated as “real”, is “wirklich”, and it has arguably somewhat different connotations than the English word “real”. As German philosophers quite often are pleased to remark, the adjective “wirklich” is obviously derived etymologically from the verb “wirken”, to act; however, the English word “real” has no similar obvious etymological connections. Schlick argued in [Sch18, §23, page 164] that only temporal things are real. Even many German philosophers whose total viewpoint is very different from Schlick’s such as phenomenologists, would agree with this. It seems likely that the word “wirklich” has a sense in which it only applies to temporal entities; however, it is not clear that the same holds of “real” and it is at least not plausible that all of the three English words would apply only to temporal entities in all of their senses. In the Section 5.5 I will deal further with questions of this kind when treating of existential multivocalism.}. Schlick tried to answer this question on the basis of verificationism, and analysed existence on the basis of the notion of regularities or law-governed connections; he came to the following conclusion in [Sch32b, page 18] (translated in [Sch49b, page 97]):

When we say of any event or object - which must be designated by a description - that it is real this means that there exists a very definite connection between perceptions or other experiences, that under certain conditions certain data appear.

This analysis permitted Schlick to say that logical positivism in its phenomenalist or neutral realist form does not deny the existence of the external world. It even led Schlick to conclude in [Sch32b, page 20] (translated into English in [Sch49b, page 98]) that the statement ”My contents of conscious-
ness exist.” is meaningless, even though the contents of consciousness are what ultimately verify all statements in his theory. Schlick tried to appeal to Russell’s symbolism in [Sch32b, page 20] (translated in [Sch49b, page 98]) in arguing for this view. He claimed that since in Russell’s symbolism an existential proposition has the form $(\exists x)f x$, the combination of words ”There is an a.”, where ”a” is a proper name of an object directly present and therefore means the same as ”this”, cannot be written in Russell’s symbolism, and hence is meaningless according to symbolic logic.

However, Schlick’s theory of existence is not compatible with his claim of the form of existential propositions, for a claim about connections and regularities would have to be formalized in Russell’s symbolism with the aid of universal quantifiers and implications and conjunctions, not with the aid of just an existential quantifier and a description. I.e. it would be of the form $(\forall x_1) \ldots (\forall x_n)(f x \rightarrow g x)$ (or a conjunction or a disjunction of such formulas), not of the form $(\exists x)f x$. This incompatibility is interesting, for it is a case where the two traditions which served as sources of inspiration for logical positivism, classical positivism and modern logic, come into conflict, suggesting a deeper incompatibility between them.

Schlick followed Russell faithfully in his view about the meaninglessness of claims of the form ”There is an a.” However, this view does not necessarily follow from the use of Russell’s symbolism, though Russell himself held it (very uncharacteristically genuinely anticipating logical positivism in this one detail). Such a combination of words where ”a” is a proper name can be written (whether a is directly present or not) in the symbolism of predicate logic as $(\exists x)(x = a)$. It is indeed taken in such a symbolization to have a more complex logical form than the syntax of natural languages would suggest, as natural languages do not suggest that such a sentence has anything to do with identity, but not to be meaningless.

2Schlick had earlier, as seen in [Sch18, §22, page 157] (translated in [Sch74, page 179]), held a more sensible view, saying that the supposition put forward by the naive view, that the directly given counts as real, must be adopted. The reason Schlick gives for this, that here is the source of the concept of reality as such, is also one that one would expect to be very compelling to a positivist. Yet strangely later logical positivists such as Schlick himself and Carnap held that reality could not be attributed to basic entities; perhaps this is the influence of structuralism derived from Neo-Kantians.

3There is indeed one qualification to be made. This formalization of existential statements in which proper names occur is inadequate in one significant way. If we use standard predicate logic (which Russell helped to create) then all such statements are logically true. This is quite outrageous, since it cannot be a logical truth that a particular (such as many
Carnap treated the concept of reality somewhat similarly in [Car67, §170-178], except that he held it to be ambiguous and therefore his theory was far more complex than Schlick’s. Carnap thought because of his verificationism that since different existential statements could be justified in different ways, they had different meanings because of this. Carnap said that he calls the only concept of reality which occurs in the empirical sciences the empirical concept of reality. According to him it is this concept which distinguishes a geographically determined mountain from a legendary or a dreamed mountain. Carnap defined empirical reality differently for physical objects, psychological objects and cultural objects. For example, physical bodies (such as the aforementioned mountains) are according to him called real if they are constructed as classes of physical points which are located on connected bundles of world-lines and are placed within the all-comprehending four-dimensional system of the space-time of physics. The empirical concept of reality is according to Carnap different from the metaphysical concept of reality, which is (see [Car67, §176, page 282]) independence from the cognizing consciousness. This concept does not according to him belong within science and is in fact meaningless⁴. Carnap also said (see [Car67, §64, page 101]) that the differentiation between real and nonreal objects does not stand at the beginning of the constructional system, which means that it cannot be applied to basic elements. We must draw especial attention to one weird feature of Carnap’s theory; it allows that physical things can be constructed as classes of physical points which are not placed within the

proper names and individual constants denote) exists. This defect in standard predicate logic helped to foster the positivists’ view that the existence of an individual cannot be stated but only shown, which in turn supported their irrationalistically mystical view that lots of important things cannot be stated but only shown. As I will argue later in more detail, in order to remove this problem we have to move from standard predicate logic to some free predicate logic.

⁴While the theory called realism maintains that independence from consciousness of whatever it is realistic about, it is doubtful if the bare word ”real” yet signifies such independence. In any case, however, Carnap’s argument for the meaninglessness of the metaphysical concept of reality is bad indeed. Carnap claims (see [Car67, §176, page 283]) that this concept cannot be constructed in an experiential constructional system, and infers that this characterizes it as a nonrational, metaphysical concept. If the concept could not be constructed in Carnap’s constructional system, this might just as well show the limitations of Carnap’s constructional system. However, while the concept may not be constructible, it can be at least partially characterized in the theory of construction. If some entities are logical constructions out of some other entities, this surely makes them dependent on those entities. Therefore Carnap’s theory does imply that physical objects are dependent on consciousness, since he claims that they can be constructed from the autopsychological.
all-comprehending four-dimensional system of the space-time of physics and such physical things are therefore unreal. In other words, Carnap agreed with Meinong that there are unreal objects.

The word "real" is surely at least on one reading synonymous with the word "existing" (as Carnap’s fellow logical empiricist Schlick testifies), so Quine would have said that real objects are those which are values of bound variables, and the concept of reality which occurs both in the empirical sciences and in ontology is the concept of being the value of a bound variable (or to express it in less semantic terms, the concept of being something). The difference of this theory from Carnap’s theory is clearly enormous. Quine would have said that what distinguishes a real mountain from a legendary or dreamed mountain is that a legendary mountain does not occur as a value of bound variable while a real mountain can. Let us for example introduce "t" as an individual constant symbolizing Mount Taniquetil (a mountain in Tolkien’s mythology) and "b" as a constant symbolizing Mont Blanc. Then the sentence (∃x)(x = t) would according to Quine be false, as there is no such entity α that for any assignment g ∥t∥g = α. However, the sentence of the same form, (∃x)(x = b), is true, as there is an such entity α, namely Mont Blanc itself, that α = ∥b∥g. This makes the difference between real and legendary or dreamed mountains clear. This concept of reality can be applied to the basic elements of a system, and indeed with the best right. Also it excludes the possibility of there being nonreal objects of any kind, even nonreal physical objects.

5While Carnap probably rejected many of the details of this specific theory later, Carnap still held later that the word "real" had different meanings when applied to different kinds of objects. E.g., in [Car56, pages 44,45] he still claimed that when we spoke about the reality of events we used the word "real" in a different sense than when speaking about the reality of e.g. electrons in general. Carnap still wanted to use such a dubious distinction to avoid traditional ontological questions. Even Carnap’s later distinction between internal and external questions in [Car50], which I will discuss in a later subsection of this work, was viewed by him as a descendent of these distinctions between different senses of the word "reality", for while it does not ascribe lexical ambiguity to the word "real" or "exist", it still claims that there are two significantly different kinds of assertions concerning reality. So this kind of distinction between different kinds of reality assertions goes through all stages of Carnap’s philosophy, so in attacking it Quine was dealing a blow to an essential part of Carnap’s type of thinking.

6There is in Carnap’s theory of nonreal objects a core that is possibly defensible, if it is separated from Carnap’s theory of reality. If a neutral monist constructs physical things from sense-data or identifies them with complexes of sense-data, then there may be constructions or complexes which are similar to physical objects but are not physical things. Carnap calls such constructions nonreal physical things, but it is better to say that they would be real but not physical; they might be called perceptual things or something
Quine’s theory is clearly far simpler than Carnap’s; one might fear that it is too simple, and as I will later show, there are qualifications that need to be made, but it still remains simpler than Carnap’s. On the other hand, Carnap’s theory brings in all kinds of considerations that are irrelevant to the question of reality, though they are important ontological considerations in their own right. Carnap may give us something that is coextensive with reality, but does not tell us what reality itself as such is. It may be true that physical points belonging to Mont Blanc are located on connected bundles of world-lines and are placed within the all-comprehending four-dimensional system of the space-time of physics. However, this does not follow from its being real.

We have seen in 3.3 that there are very strong (even if not wholly conclusive) reasons to think that verificationism is false, and this gives us a reason to reject Schlick’s and Carnap’s analyses of reality and existence as well as the causal analysis of reality and existence. As verificationism forced him to do, Schlick confused the question of how we can justify the claim that something exists or is real, and the question of what that claim means. Quine’s criterion of ontological commitment avoids this confusion (though I have argued that Quine falls into the very same kind of confusion elsewhere in his philosophy, in his arguments against the distinction between analytic and synthetic judgements). Quine’s theory, unlike Schlick’s, is really based on the philosophical application of modern symbolic logic and Russell’s symbolism (and can indeed be viewed as being little more than a return to the view of existence common in analytical philosophy before the rise of logical such. Of course, this theory is false if physical things are not constructions out of sense-data - to say nothing of what is the case if there are no sense-data at all - but it is compatible with Quine’s theory of reality.
positivism\textsuperscript{7}, with some clarifications and simplifications\textsuperscript{8}).

5.2.2 Bergmann’s Theory of Ontological Commitment

There is yet one theory of ontological commitment that is influenced by logical positivism, though it is already partly liberated from logical positivism, namely Gustav Bergmann’s theory. Bergmann said in [Ber54, page 135] of classical ontologists:

When these philosophers asked themselves what existed, they really asked for the entities that were named by their undefined descriptive terms.

Bergmann calls descriptive terms what are currently more commonly called non-logical constants (of an artificial language, not any natural language). Expressed in the terminology I have been using, this implies that according to Bergmann, a theory carries ontological commitment to entities which are denoted by the primitive non-logical constants of the language the theory is

\textsuperscript{7}Quine’s view is similar to the view of the common concept of reality G. E. Moore expressed in [Moo22a, page 211] when he said that to say “Unicorns are unreal,” means the same as ”There are no unicorns,” or ”Unicorns do not exist. Moore then almost explicitly denied Carnap’s later assumption that there are unreal objects. Moore’s view is already incompatible with the view of such current opponents of Quine’s theory of ontological commitment as Jody Azzouni, who denies in [Azz04] that ”there is” expresses existence. This is rather ironical, since both Moore and Azzouni claim to be defenders of common sense. Criticizing Bradley’s idealistic theory, Moore did suggest that Bradley might have used the word ”real” in a special sense. This would fit with Carnap’s distinction between empirical and metaphysical concepts of reality (though you could not really infer from Bradley using the word ”real” in a peculiar sense that all or most metaphysicians would have used it in a peculiar sense, as Carnap claimed). However, Moore ultimately concluded that such an interpretation would not solve the many problems in interpreting Bradley’s obscure theory. A difference between Moore and Quine is of course that Moore did not connect his theory of existence with the quantifiers and variables of artificial languages, since he was not an artificial language philosopher as Quine was. Later in [Moo53], Moore had a view which is similar, but not quite as close to Quine’s later one. Moore accepted in that book a weak version of existential multivocalism, saying in [Moo53, chapter XVII] that some things (namely fact and universals) probably are without existing.

\textsuperscript{8}Among such simplifications are the following: there is no evidence that Quine would have accepted Frege’s and Russell’s idea of existence as a higher-order property, though such a conception has sometimes been ascribed to him (e. g. apparently by William F. Vallicella in [Val03]). Quine seems rather to consider the quantifiers of predicate logic (and hence such natural language expressions as ”existence”) themselves as a syncategorematic expression. Nevertheless, Quine’s theory could rather easily be modified so that it would take existence as a higher-order property, so Vallicella is right so far that Quine’s theory has similarities to the higher order theory of existence. Another simplification is that Quine does not accept Russell’s distinction between existence and subsistence (or Moore’s similar distinction between being and existence).
formulated in. These undefined descriptive terms could be either individual terms or predicate terms.

One very weird thing about this theory is that according to it the ontological commitments of a theory depend only on the language of the theory, not at all on what statements are contained in that theory. This is surely already a reason to be suspicious of Bergmann’s theory. There are other reasons. The theory presupposes strong fundamentalism, since it implies that only simple entities exist in the ontological sense. I have already argued that we have no good reasons to believe in strong fundamentalism, so that would already be a reason to reject Bergmann’s theory. However, it seems that Bergmann’s theory also presupposes verificationism, so that it is unacceptable also to a fundamentalist who is not a verificationist. Primitive descriptive terms are just the terms associated with entities we are acquainted with (unless they are non-denoting i.e. empty); they are observational terms in the language of later logical positivism. Bergmann’s theory then implies that ontology can only contain entities we can observe. Bergmann’s theory implicitly denies the possibility of knowledge by description. Bergmann did not see that even a simple entity may be denoted by a complex description, if the description picks it out by means of its relational properties instead of its internal constituents.

Summing up our results, we can safely say that Quine’s theory of existence or being or reality, then, is clearly an improvement over such positivistic theories as Schlick’s, Carnap’s and Bergmann’s, and even the Eleatic Principle so far as it competes with it. However, this does not yet show that it would be wholly correct. Was Quine right or wrong? Unfortunately, the question has no simple answer. Quine formulated his criterion in different ways in different places. Quine apparently thought that all these formulations were equivalent; however, this can be doubted and I will show that all of them are not equivalent. Thus even if some of them are correct all of them cannot be. Trying to reconcile all of Quine’s often slapdash utterances into one systematic theory in the spirit of authoritarian respect, in the way scholastic philosophers treated Aristotle’s texts, would not be a promising way to set about constructing a theory of ontological commitment. Any viable interpretation of Quine’s theory must be mixed with a healthy amount of criticism. Besides this most of Quine’s formulations are not wholly free of obscurity. Therefore it is possible that most of his formulations will be
correct under one legitimate reading but incorrect under another. I will argue that this is in fact the case.

5.3 The Eleatic Principle

There have been many other theories of existence which can also be argued to be based at least partly on verificationism (even if their authors do not always recognize this). One very popular theory is that the criterion of existence is (ability to exert) causal influence. David Armstrong used such a criterion in [Arm78] to argue against uninstantiated universals. This criterion was dubbed by Graham Oddie (who did not find the principle correct) in [Odd82] the Eleatic Principle, after the Eleatic stranger who in Plato’s dialogue Sophist (247e) propounds a similar principle. This name has become popular; Armstrong himself used it in [Arm97, §3.82, pages 41-43]. It has also been called by other names such as Alexander’s dictum and the causal criterion (of ontological commitment). This principle seems to be the most popular current rival to Quine’s criterion.

The principle has been presented in various formulations, which need not be equivalent, as Oddie already pointed out. It could be understood as either a logical or a factual principle. Armstrong prefers (see [Arm97, §3.82, page 41]) a relatively weak formulation according to which everything which exists must make a difference to the causal powers of something instead of a stronger formulation according to which everything that exists must itself have causal power. Similarly Oddie distinguished between a stronger formulation that appeals to being causally active and a weaker formulation that appeals to playing a role in the causal structure of the world and Colyvan distinguishes between being causally efficacious and being causally relevant. I would guess that all three distinctions are equivalent.

The most important argument for this theory is that if something was incapable of exerting causal influence on us we could not have any evidence that it existed. This argument is often directed specifically against the existence of mathematical entities such as sets, and therefore against the literal truth of mathematical theories, following a line of thought presented by Paul Benacerraf in [Ben64] (though Benacerraf did not yet use the name ”Eleatic Principle). Benacerraf himself did not yet draw the conclusion that mathematical entities did not exist, for he presented a dilemma to which he did
not see any satisfactory solution. However, many later philosophers have not hesitated to grasp the horn of the dilemma according to which mathematical entities do not exist and mathematical statements are not literally true. However, there has not been unanimous agreement that mathematical entities would be incapable of exercising causal influence or of making a difference to the causal powers of something. E. g. Penelope Maddy has argued in [Mad80, page 182] that sets can be causally responsible for perceptual beliefs\(^9\). Oddie also gives in [Odd82, page 293] an example of how mathematical objects could play a role in the causal structure of the world, which does not appeal to perception. Suppose that Jones weighs 100 kilograms. As a result whenever he climbs the stairs in his house they creak. His being one hundred kilos certainly plays a causal role in this type of event and so it would seem that derivatively the number 100 also plays a role.

The Eleatic Principle is even often used to argue against all kinds of realism about universals, which is rather ironic since the name was originally coined to apply to a principle by David Armstrong, who was an immanent realist. The Eleatic Principle works at most (as Armstrong intended it to work) against transcendent or "Platonist" realism, that takes universals to be outside space-time, and not\(^10\) immanent or "Aristotelian" realism, which holds universals to be multiply located within space-time\(^11\).

\(^9\)Maddy changed later her views in the philosophy of mathematics, and opted for a less realistic account of mathematics. However, she does not appear to have backed down from the view that accepting the existence of mathematical entities is compatible with a causal theory of knowledge.

\(^10\)A lot of discussions in the philosophy of mathematics - as opposed to discussions in general ontology - contrast "Platonism" with Nominalism as if these exhausted the options and just leave "Aristotelianism" out of consideration. If such discussions recognize the possibility of an "Aristotelian" alternative at all, it is only as a belated afterthought, and they usually try to assimilate it to one of the two options than treat it as an independent type of theory. Thus Mark Balaguer argues in [Bal98, §2.5.1, pages 29-31] that Maddy’s theory is really platonistic and that she does not really view mathematical objects as concrete as she claim, but must treat them as abstract in some non-traditional sense. Balaguer’s argument is that if Maddy were really non-platonistic, she would have to admit that sets are aggregates of physical matter. However, it is by no means obvious that aggregates (mereological sums?) of physical matter are all that is spatio-temporal. Weirdly enough, Balaguer’s own position seems to be "Aristotelian", though he says it is nominalistic, for he says in [Bal98, §6.5, page 127] that physical properties are nominalistically kosher. They are certainly not such if the word "nominalistic" is used in its common sense. If Balaguer can admit the existence of physical properties, why not say that mathematical objects are physical properties? This would indeed differ a bit from Maddy’s theory, but be but a small modification of it.

\(^11\)I use the scare quotes because (as I show in Section 6.3 of this dissertation) it is highly doubtful if these two kinds of theories are genuinely similar to Plato’s or Aristotle’s
The Eleatic Principle is faced with many problems, especially in its stronger versions, many of which Oddie already pointed out when providing it with a name and many more of which Mark Colyvan points out in [Col98]. There are of course metaphysical theories, such as Bertrand Russell’s view in [Rus12a] that there are no causes, which would imply together with the Eleatic Principle that nothing exists, which would be absurd. However, even if Russell’s arguments for his view are not considered sufficient, problems remain.

Stronger formulations of the principle are also in danger of collapsing into forms of verificationism; Armstrong himself did not want them to do so, which caused him to leave his discussion of the Eleatic Principle in [Arm97, §3.82, page 42] in uncertainty. However, Evan Fales defends in [Fal90, page 213] a doctrine he actually calls neo-verificationism which as an ontological principle says that something which (causally) cannot make a difference to the rest of the world (hence to our senses) has no physical existence at all. On the other hand, weaker formulations of the principle threaten to collapse into vacuity. Oddie suggests in [Odd82, page 2929] a formulation according to which nothing is real unless it is at least conceivable (intelligible) that it should play some role in the causal structure of the world. However, such a formulation would contrary to Armstrong’s intentions allow even the existence of uninstantiated universals, for if a universal is un instantiated but it is conceivable that it should be instantiated, and if it would play a causal role when instantiated, then it is conceivable that it should play a causal role.

The argument for the Eleatic Principle presented above of course presupposes a causal theory of knowledge or justification (such as was sketched by historical theories.

12Some of Russell’s arguments presuppose strong fundamentalism, which I have argued against. The argument that the word ”cause” never occurs in advanced sciences only shows, if we do not accept strong fundamentalism, that causation is not fundamental, not that it does not exist. However, even this conclusion, if correct, would be enough to refute the Eleatic Principle, for surely more fundamental entities exist if less fundamental entities do.

13Mark Balaguer argues in [Bal98, §2.2, page 23] that Benacerral’s epistemological argument against (traditional forms of) platonism can be reformulated so that it does not depend on a causal theory of knowledge. He suggests that the premise of the argument can be reformulated so that they say that mathematical objects are totally inaccessible to us, i.e. that information cannot pass from mathematical objects to human beings. This gives according to him rise to a worry about whether human beings could acquire knowledge about mathematical objects. However, it is not at first sight at all clear what Balaguer
Paul Benacerraf in [Ben64, page 671]), which is quite controversial, though popular\textsuperscript{14}. If some such theory were correct, then it may be correct that if something was completely causally isolated from us, we could not have any reason to think that it existed. However, this would scarcely give us a reason to think that no such entity exists either. Rather, the best response to a claim about the existence of such entities would be to suspend means with information and its passing. Information itself, as commonly understood, is something abstract, so it could be argued that if we can have information at all about any objects, we must already be related in some way to an abstract object. Balaguer says later in [Bal98, page 25] while arguing against Gödel’s view that mathematical objects could not generate information-carrying signals, which makes the argument a bit clearer. However, it is not clear that knowledge presupposes that information-carrying signals should pass from the objects of the knowledge to the knower any more than that there is a causal relation between them. Balaguer adds later in [Bal98, page 26] that the notion of an information transfer is a causal, spatio-temporal one, and this indeed clarifies matters further. However, it seems also to reduce his purported reformulation of Benacerraf’s argument against traditional forms of Platonism to its original formulation. If according to Balaguer the acquisition of knowledge requires information transfer (from the objects of knowledge to the knower) and information transfer is a causal relation, then it follows that according to him knowledge requires a causal relation between the knower and the objects of knowledge. Balaguer is therefore presupposing a causal theory of knowledge after all in his argument. This kind of theory seems to share all of the same problems as a causal theory of knowledge, which is natural if it is after all just one form of it. E. g. it would make knowledge of the future impossible. Could information-carrying signals pass from future to the present either? What difficulties the "Platonist" has in accounting for mathematical knowledge are then not greater, so far as the argument goes, than the difficulties anyone has in accounting for knowledge of the future.

\textsuperscript{14}The eleatic principle is not directed against the possibility of ontology as the original verificationist principle was but only against specific ontological theories. However, it is similar to the verificationist principle in one respect; it is based on an epistemological principle that is mostly appealed to in discussions concerning ontology and only seldom held to be plausible in epistemological discussions which have nothing to do with ontology. While causal theories of knowledge have occasionally been developed by philosophers specialising in epistemology, these theories have been criticised heavily as epistemological theories and most epistemologists have shifted away from them to other kinds of theories. For example Alvin Goldman argued for a causal theory of knowledge in [Gol67], to which Benacerraf indeed referred, but later (e. g in [Gol86]) he replaced it with a more general reliabilist theory of justification that cannot be so easily be used to justify the Eleatic Principle. It is easy to see what serious problems there are in the principle in question if it is held in an unqualified form. This principle would make impossible any knowledge of the future (or more generally as Colyvan notes of entities outside our light cone), since future entities cannot affect anyone’s present state of knowledge. Thus it would make all knowledge useless, since knowledge is only practically useful if it can be used to predict the future. Of course attempts can be made to weaken the causal theory as attempts were made to weaken the verificationist principle. The causal theory can be weakened so that it only demands that epistemically basic knowledge is of entities affecting the knower, while less basic knowledge can be derived from this basic knowledge inductively or abductively. However, so weakened the principle might allow the existence of even transcendent universals, since induction or abduction might then be used to support belief in such universals.
judgement about their existence\textsuperscript{15}. Furthermore something could make a
difference to the rest of the world without making a difference to our senses,
and such an entity would also be unknowable to us. Therefore this kind of
argument would if valid at all lead to a far stronger conclusion, namely that
the criterion of existence is ability to exert causal influence on us. However,
this sounds (as Mark Colyvan also points out in [Col98]) rather too
anthropocentric to be credible.

\section*{5.4 Motivating Quine’s Theory}

Quine does not really motivate his suggestion at length, intuitively appealing
as it is. A more elaborate motivation is clearly needed. Unfortunately, I can
for reasons of space only sketch such a motivation in this dissertation, which
will rather concentrate on the interpretation and consequences of Quine’s
widely accepted suggestion than its motivation.

It must be admitted that it is by no means obvious at first sight that
the suggestion is true. At least linguistically quantification is expressed in
natural languages very differently than being; quantification is expressed
primarily with the aid of determiners such as ”some”, ”every”, ”most”, etc.
while being is expressed primarily with the aid of verbs like ”be” and of
course ”exist” together with prepositions like ”there” and adjectives like
”real”. It has often been argued with some cogency that what is usually
called the existential quantifier would be more accurately called the particu-
lar quantifier. However, even long before the rise of modern logic it was
rather easy to see that natural language sentences formulated with the aid
of quantifiers and sentences formulated with the aid of existential verbs were
often logically equivalent. Consider the following example sentences:

\textsuperscript{15}Even those who think that one has to be non-realistic about entities with no causal
powers need not deny that theories, even scientific theories, could carry ontological com-
mitments to such entities. Rather they could treat causally inert entities as van Fraassen
does unobservable entities; namely recognize that a theory can meaningfully carry com-
mitments to them but hold that such theories cannot be known to be true and treat
such portions of a theory which make such commitments instrumentally. In other
words, using Quine’s terminology (see Section 5.8), the Eleatic principle would accord-
ting to them afford a criterion for adjudicating between rival ontologies, not for recognizing
the ontological commitments of a discourse. Indeed, many proponents of the principle
seem to use it as such a principle for adjudicating between different ontologies; Armstrong
says explicitly in [Arm97, §3.82, page 42] that it does not have to be taken as a semantic
principle. However, the arguments of Oddie and Colyvan are effective also against the
principle if taken as a criterion for adjudicating between different ontologies.
Example 1  *There is a mortal man.*

Example 2  *A mortal man exists.*

Example 3  *Some man is mortal.*

The sentences Example 1 and Example 2, in which being or existence is explicitly expressed, are clearly equivalent with the sentence Example 3, in which only a quantifier occurs, at least given one very natural and common reading of these three sentences. Obviously the equivalences would remain valid if the noun phrase ”mortal man” and the adjective ”mortal” were in all these sentences replaced systematically with almost any phrases of the same syntactic and semantic category, with only some rare exceptions like ”alleged criminal” and ”is fictional”, so there are here an infinite number of potential examples. Such equivalences clearly suggest that the particular quantifier can also be legitimately called an existential quantifier. Modern logic and semantics show systematically what such equivalences hold and to some extent why they hold; furthermore they suggest that the existential quantifier is a basic quantifier with whose aid all other quantifiers can be defined, using no other quantifiers but many other other notions besides it (including set-theoretical or property-theoretical notions in the case of such semantically complex quantifiers as ”most”). However, non-standard logicians have suggested some purported counterexamples to such equivalences; I will deal briefly with one prominent counterexample later.

Drawing on such intuitive equivalences between quantification and existence, Quine’s suggestion can be motivated if we can also show that there is a close connection between the concept of ontology and the concept of being; however, I have already presented historical evidence for such a connection in Chapter 2 of this work.

5.5  **Existential Univocalism and Multivocalism**

Matters are complicated by claims that being is not univocal. With Quinean metaontologists the criterion of ontological commitment is usually closely connected with the thesis of the univocity of being. This thesis has been called univocalism by Berti in [Ber01] following Morton White (who also
defended it in [Whi56, page 61] and existential monism, as the phenomenological ontologist Roman Ingarden, (who opposed the thesis) called it in [Ing64, page 48]. For instance, Peter van Inwagen explicitly and strongly defends the univocity of being in [vI98, page 236]. White called the opposite of univocalism multivocalism while Ingarden spoke of existential pluralism. Many philosophers such as Putnam (see [Put94, page 450]) and Eli Hirsch following him (see [Hir02]) have objected to Quine’s criterion exactly because it presupposes univocalism.

However, denials of the univocity of being are far older. Aristotle himself already thought that being was not a genus and that there were many different senses of being. For instance, as I will show in Section 4.2.3, Aristotle thought that states of affairs existed in a different sense than substances or even accidents, and hence were not full-fledged members of his ontology. These different senses of being Aristotle distinguished were according to him not entirely heterogeneous but connected, the relation between them being what medieval philosophers called analogy (more specifically analogy of proportion in Cajetan’s terminology) and what is sometimes today called core-dependent homonymy. Aristotle gave the following kind of examples to illuminate this notion: a man can be called healthy and food can be called healthy in a different sense, but the second sense can be defined with the aid of the first, for food is healthy if eating it helps a man to keep healthy; Aristot-

\[16\] White claims in [Whi56, page 62] that John Stuart Mill also earlier defended univocalism but also expressed a number of worries over it that have reappeared in more recent literature. Looking at the places in Mill’s writings that White refers to I cannot say that I am convinced by White’s interpretation. Rather than defending univocalism, Mill seems to rather unquestioningly presuppose it, and asks what term we should use for that which exists in this one sense of the word “exist”: the terminological questions that Mill deals with are far more trivial than the genuinely semantical and even ontological questions discussed in later literature. These terminological questions are, however, still relevant; I use the word “entity”, but other ontologists have other uses.

\[17\] Apparently Aristotle and medieval philosophers thought that being was not one of Porphyry’s other five predicables either. The Arabic philosopher, Ibn Sīnā or Avicenna, apparently challenged this view and held very interestingly that existence was an accident of essence. The distinction between existence and essence that Avicenna made (with some predecessors who were not so clear about the matter going back to some obscure terminology and remarks of Aristotle himself) was also a kind of distinction between different sense of being. However, Avicenna’s view did not become popular either in Arabic or western scholastic philosophy.

\[18\] How many modes of being Aristotle distinguished is a bit obscure and controversial, since Aristotle divides being differently in different places and it is not clear how (if at all) these divisions fit together. Gareth B. Matthews examines this question in [Mat95], showing that there are reasons to think that Aristotle distinguished either as many as ten or as many as fourteen or as many as forty or even indefinitely many.
tle claimed that the different senses of being were related in this way. Very startlingly, however, as Enrico Berti remarks in [Ber02], Aristotle almost never brings any argument in defence of this doctrine; actually Aristotle has only one passage that can be viewed as proof of this doctrine (and this argument is very weak since it is based on his very narrow theory of definition by means of genus and differentia). As Berti says, by so doing Aristotle gives the impression of considering his doctrine perfectly evident, though he was clearly persuaded that he was the first philosopher who discovered this truth. Nevertheless this became a standard dogma in medieval scholasticism; because of the authoritarian nature of most medieval philosophy, scholars accepted Aristotle’s pronouncements meekly as the truth. Notions such as being and unity which were not genera or species but could in some way be predicated of all entities were called transcendental by the scholastics; due to their ridiculously great theistic optimism they also thought that good could also be predicated of all entities and was also such a notion. Medieval philosophers expanded Aristotle’s claim to the theological claim that God existed in a different sense than creatures, thereby making it a religious dogma and ensuring it would not be questioned. However, already some medieval Aristotelians such as Duns Scotus at least partially denied this assumption of Aristotle himself and claimed that there was a concept of being univocally common not only to entities of different categories but

---

19The argument occurs at *Metaphysics* 998b22-27. It starts from the premise that a genus cannot be predicated of its differences (or differentiae, however we translate διαϕορα) by means of which species are split off from genera, and shows that if being were a genus then its differences (differentiae) could be predicated of it, since being would have to be a genus of all entities. The argument is formally valid, but its premise, that a genus cannot ever be predicated of any of its differences, seems completely unjustified to me. Indeed, the whole distinction between species and differences seems very obscure and dubious from the point of view of modern logic and semantics. So far as I can make any sense of the distinction, it is based on the notion of identity conditions: species and genera are supposed to provide a identity conditions for the entities of which they are predicated, but differences are not. However, as I point out also elsewhere in this work, this whole notion of identity conditions has been questioned by such philosophers as Michael Jubien in [Jub96].
Of course, many later philosophers have also claimed that being is not univocal, but have divided its various senses in very different ways than Aris-totelians did. This kind of doctrine can be found in all major modern philo-
sophical traditions, as well in analytical philosophy (both non-positivist and positivistic) as in phenomenology.

Famously Frege and Russell claimed that the word ”being” (or concretely ”be”) could stand for existence, predication, class inclusion and identity; we have already seen how this theory has been opposed on the basis of Charles Kahn’s historical studies. This is a far smaller variety of modes of being than the one Aristotle arrives at. Furthermore, it is one that Quine himself and van Inwagen probably would have accepted. When van Inwagen says that being does not differ from existence, he carefully explains in [vI98, page 236] that what he means to deny is that the phrases ”there is” and ”exists” mean the same thing when applied to different subjects. This does not imply that he would have to deny that being as predication differs from being as existence. What neither Quine nor van Inwagen would not have accepted is the claim that ”existence” or in more concrete terms such phrases as ”there is” and ”exists” (which are according to Quine symbolized by the existential quantifier) or indeed ”some” have themselves many senses or uses.

Lambertus Marie de Rijk, who rejects Kahn’s theory of the development of the Greek verb yet follows him in rejecting the Frege-Russell thesis of ambiguity, says in [Rij02, §1.4, page 24]:

...contemporary analytical philosophy is generally inclined to emphasize that once the various notions involved in our use of ’be’ have been logically analysed, they appear to form a heterogeneous bundle without any focal concept of ’be’ to hold them together. It will be plain that to establish an unambiguous sys-
tem of metaphysics must seem a chimaeric endeavour to these logicians.

However, there are reasons to fear that Scotus’s denial of the univocity thesis was very partial; Scotus apparently insisted that the application of a common concept to different entities did not imply that they shared a common nature, and because of this his denial did not have metaphysical but only semantical implications (see for example [Dum92]. Even though Aristotle did not give almost any arguments for his claim that being was equivocal, the authoritarian nature of mediaeval philosophy kept medieval philosophers from questioning it seriously, so that rejecting it completely would have been a serious option for them.
However, the modern logical theory by no means implies that the notions of being would be a heterogeneous bundle, though this is a common misunderstanding. Many of these notions can be defined with the aid of each other and even in the case where such definition is not possible, a sentence containing one of them always implies a sentence containing others, so these notions are very tightly connected together. Class inclusion can of course be defined with the aid of quantification, implication and membership (or just predication in higher-order logic); $A \subseteq B \equiv df (\forall x)(x \in A \rightarrow x \in B)$ in first order logic combined with set theory or $A \subseteq B \equiv df (\forall x)(A(x) \rightarrow B(x))$ in higher-order logic. Identity can be probably defined with the aid of universal quantification and predication in higher order logic or set theory; $a = b \equiv df (\forall P)(P(a) \equiv P(b))$ or $a = b \equiv df (\forall x)(a \in x \equiv b \in x)$, though this is more controversial; in any case, saying that two entities are identical always involves predicating identity of them, so predication always goes together with identity and is therefore more fundamental even if identity could not be reduced to it. Therefore while there is no single focal meaning, yet existence and predication are two focal meanings on which the rest of the notions are based, so the modern theory can be viewed as a version of an analogical theory of being, even though the principle of the analogy in question is more complex than in Aristotelian logic (though simpler in its applications). Furthermore, these two always go together. You cannot formulate a statement in the language of predicate logic (at least not in a standard first-order one) that contains a quantifier but no instance of predication or identity. Neither $(\exists x)(x)$ nor $(\exists x)(P)$ is well-formed; the simplest well-formed quantified formula is of the form $(\exists x)(P(t))$, and the more complex formulas always contain one or more instances of predication. A predicative statement $P(a)$ implies in standard predicate logic and negative free logic such quantified existential statements as $(\exists x)(x = a)$, $(\exists x)(P(x))$, etc. The notions of being distinguished by Frege and Russell are then very tightly bound together so establishing an unambiguous system of metaphysics is a quite tenable goal on the basis of their theory.

However, Russell held also the kind of serious existential multivocalism that Quine and van Inwagen oppose in a weak form, as he found it convenient in [Rus12b, pages 155-156] to distinguish existence from subsistence. According to him things exist only when they are in time, while timeless entities (which universals according to him were) subsist or have being. Russell
apparently later rejected this distinction. Russell also developed the theory of types as a solution to the set-theoretic paradoxes, and it is often held that the theory of types supports multivocalism. According to it there are as many types of quantifiers as there are types, i.e. infinitely many, and if existence is expressed by quantifiers, then we can infer that a plurality of quantifiers implies a plurality of modes of existence. It is not clear if Russell himself would have accepted such a corollary to his theory. In any case, since later set theory and property theory have mostly rejected type theory, this naturally weakens the force of this reason for existential multivocalism.

Russell’s notion of subsistence comes from the German (or perhaps rather Austrian) tradition; Meinong had already spoken of subsistence (bestehen) as the mode of being of ideal entities such as universals or propositions (which must be sharply and carefully distinguished from Meinong’s far more controversial impossible objects). Many phenomenologists made a similar distinction speaking of the ideal being of universals.

However, the Polish realist phenomenologist, Roman Ingarden, was unique in developing a very elaborate and complex theory of different modes of being in his main work *The Controversy over the Existence of the World*, originally written in Polish, of which only a very small amount has been translated into English in [Ing64] but far more of which has been translated into German. Ingarden distinguished in [Ing64, page 22] a basic group of ontological questions, existential-ontological questions, which were concerned with modes of being. Ingarden distinguished in [Ing64, page 41] moments of existence aka existential moments in modes of being, of which he in a way built up these modes of being. Unfortunately Ingarden’s theory cannot be discussed here in any detail for reasons of space.

Logical positivists also sometimes denied univocalism; Ayer claimed on the basis of his neutral monism in [Aye36a, page 130] that

\begin{quote}
when one says that a sense-experience or sense-content exists, one is making a different type of statement than when one says that a material thing exists.
\end{quote}

Ayer took entities which he supposed to be logical constructions (such as material things and minds) to exist in a different sense than entities which he took to be primitive. Carnap says similarly in the *Aufbau* (see [Car67, §42, page 70]) that one could speak of the different modes of being of the
objects of different object spheres (where object spheres are the levels of logical construction), and gives as an example (see [Car67, page 71] that classes do not have being in the same sense as the things from which they are in his view constructed\textsuperscript{21}. This really amounts to claiming that quantification over logical constructions is non-objectual; this claim is not peculiar to Ayer but is a view which is implied by strong fundamentalism and which all strong fundamentalists implicitly hold.

As White and van Inwagen show, even ordinary language philosophers such as Gilbert Ryle and Austin held that being was equivocal and were influenced by Aristotle in doing so. They actually tried to use this to argue against metaphysics, paradoxically justifying their opposition to metaphysics by Aristotle’s metaphysical doctrine. Ryle for example claimed in [Ryl49, page 23] that when we say bodies exist and when we say minds exist we use the word ”exist” in different senses, and because of this we could not say that both minds and bodies exist If we accept Quine’s claim of the connection between existence and quantification, then the acceptance of different kinds of quantification counts as the acceptance of different kinds of existence. Therefore if we must distinguish from each other objectual and substitutional quantification, this is already a weak kind of descriptive existential pluralism.

In most of the forms we have gone through, multivocalism is not connected with any kind of relativism. Most multivocalists (e. g. Aristotle and Ingarden) thought that expressions like ”is” or ”exists” have many senses in the same ordinary language, and that the same metaphysical theory must distinguish several modes of existence. This allowed several different metaphysical theories to speak about the same modes of existence. It seems to me that this still held also of Carnap’s theory in the \textit{Aufbau}. However, with the last stages of logical positivism and ordinary language philosophy emerged a relativistic form of multivocalism that is still prominent in metaontology. This shift to relativism resulted from a combination of multivocalism with

\textsuperscript{21}Logical positivists did not agree on this point. Though Schlick is one of the major sources of Ayer’s doctrines, on this point Schlick seems to have differed from Ayer in [Sch32b, page 20] (translated in [Sch49b, page 98]), denying that the claim that a feeling or sensation is real would have a meaning different from the claim that a physical object is real. However, as I show in comparing Schlick’s theory of existence and reality to Quine’s theory in 5.2.1, Schlick’s theory is utterly incoherent and in fact I think that Ayer’s theory is a more coherent application of the verificationism both shared to the question of existence, though of course also suspect to those who reject verificationism.
semantic holism based on verificationism. Eli Hirsch claimed in [Hir02] on the basis of the ideas of Hilary Putnam that when philosophers apparently disagreed with each other or with common conceptions on what existed, they were just using different concepts of existence so that they did not really disagree. On Hirsch’s view then different metaphysical theories always employed different concepts of existence. Matti Eklund calls this kind of view ontological pluralism and claims in [Ekl09] that Carnap held such a view in [Car50]; I will examine this kind of relativistic view in detail in Section 5.9.2 of this work and argue that it is not tenable.

If being is really equivocal it becomes unclear what the definition of ontology as the science of being qua being means. Should we allow that there are several ontologies corresponding to different senses of ”being” aka different modes or ways of being, or should we pick one of them as the one true ontology? How are these possible different senses of being to be distinguished from each other anyway?

I propose that contrary to what Quine himself and van Inwagen may have thought, the criterion could be modified so that it does not presuppose the univocity of being, at least if the univocity thesis is understood as a descriptive thesis concerning natural languages. What happens to the criterion if words like ”be” or even ”exist”, as they are used in ordinary languages, are equivocal? Does this show that the criterion is false? No, it only follows from this that the notion of ontological commitment is also equivocal. However, I think that this would be a quite acceptable result. As we have seen, it was held by the very founder of metaphysics, Aristotle. Many modern ontologists such as Ingarden have held that there are many modes of being, and if the word ”be” is ambiguous then this view of existential pluralism follows from Quine’s criterion.

In fact, it seems to me that there is such ambiguity in natural language. Quine’s claim of univocity in fact fits in very badly with his own later thesis of the indeterminacy of denotation, and even though there are good reasons to think that denotation cannot be quite as indeterminate as the later Quine himself claims (for in that case we could not talk intelligibly of anything at all), yet there is enough indeterminacy in natural languages and whatever artificial languages have been constructed that the absolute univocity of any concept is as such very implausible. Thus we should except to find
multivocity, and in fact there is detailed evidence for such multivocity. Wolterstorff has in my view argued very convincingly in \[\text{Wol70, page 167}\] that the words "there are" are ambiguous when applied to universals. Saying "there’s no such thing as genuine humility" can naturally be used to mean that humility is not exemplified. On the other hand, the sentence "There’s no such property as being a round square." has a stronger meaning as it implies that there could not be round squares. This variety of uses could, of course, be also interpreted as a purely pragmatic matter. Even if it is a genuinely semantic phenomenon, however, it does not really threaten the criterion of ontological commitment. However, it does raise the question what this implies for the unity of ontology as as scholarly discipline.

However, the thesis of the univocity of being is better understood as a partly revisionary than as wholly a descriptive proposal (as Owen Pikkert also argues in \[\text{Pik12, page 7}\]). Even though all quantification in ordinary languages would be ambiguous, this does not imply that the ideal languages or even artificial languages we can in practice construct for purposes of scholarly research should be ambiguous in a similar way. Ambiguity should of course always be eliminated in philosophical systematization (what Quine calls regimentation) of common sense notions so far as it can be eliminated. Putnam and even Quine may, of course, think because of their semantic holism that ambiguity cannot ever be eliminated; however, as I have already argued this kind of holism is in itself quite implausible. Nevertheless, we can only know how far we can eliminate ambiguity by trying and seeing what is the result.

However, eliminating ambiguity may involve symbolizing the same expression of a natural language with several primitive expressions, or using only one expression to indicate one of the senses of the natural language expression and introducing the rest of the senses as defined expressions. The debate between univocalists and multivocalists can be understood as involving which of these two options should (or could) be adopted; Quine’s view would be that it is the latter option. It then implies a position that is the very opposite of the one Hirsch maintains; Hirsch thinks that being is not equivocal in ordinary languages, but claims that when different philosophers

\[\text{Footnote 1 in the previous section.}\]
try to regiment it the results of their work differ from each other. However, the way quantification is actually ordinarily used may not be the most useful one! This is something that opponents of Quine’s criterion like Putnam or Hirsch do not seem to realize. Just distinguishing different senses of quantification is not enough; we must ask about how these senses are related. One question concerning their relationship is paramonally important. Given that there are many modes of being, we must ask whether they can be in some way reduced to one mode of being.

Many philosophers have thought that they can. Aristotle himself thought that there was one sense of being, that in which actual substances (whether concrete composite individual substances or forms are in question is disputed) existed, which was the primary sense of being and the only strict and proper sense of being, in terms of which the other senses of being were to be characterized. The same holds of most later theories that try to divide the senses in which entities exist; they may not agree that substances exist in a primary sense (or exist at all) or even that being is not a genus, but most of them agree that there is one basic sense or there are a few basic senses into which all of the remaining ways of being have to be related (which according to some of them is indeed a genus under which all or at least some of the remaining ways of being fall). If this is possible, then many kinds of quantification can be replaced with one basic kind of quantification in artificial languages. I will later in Section 5.9 argue that objectual quantification is primary with respect to other kinds of quantification such as substitutional. However, the argument is far from conclusive. Even I were wrong in this and there were many primary kinds of quantification, this need not yet lead to any disastrous consequences for ontology such as relativism, since as we have seen most forms of existential multivocalism are not relativistic.

I think that there are two basic senses of ”being”, predication and absolute real existence (where the latter is the sense that can be symbolized by an unrestricted objectual quantifier). These senses cannot be reduced to each other, but they always go together, for when we say that entities of some kind exist, we always predicate something, in the simplest cases that kind itself, of some variable bound by the unrestricted objectual quantifier symbolizing absolute real existence. I will show that absolute real existence can be viewed as a genus under which many other forms of being (such as restricted quantifiers) fall, but many other forms of being have to be related
to it in a more complex way, closer to the Aristotelian way of analogy.

5.6 Unrestricted Quantification and Absolute Generality

We have seen that ontological statements in the sense of Wolff and metaphysical statements in the sense of Aristotle’s conception of metaphysics as the science of being as being are absolutely general; they concern all entities, entities belonging to any domains. This means that they employ unrestricted quantification\(^{23}\). However, the meaningfulness of unrestricted quantification has been disputed by many philosophers, which threatens the possibility of ontology or metaphysics in this historical sense. However, lately it has gained many defenders. I think that the most important of its defenders is Richard Cartwright in \([\text{Car94}]\). Other more recent defenders of absolute generality have been Timothy Williamson in \([\text{Wil03}]\), Øystein Linnebo, Vann McGee and Alan Weir. Both Cartwright and Williamson point out that commitment to unrestricted quantification is essential to Quine’s philosophy. Cartwright points out in \([\text{Car94}, \text{page 1}]\) that variables are to be regarded as taking as values any objects whatever in Quine’s set theory NF. Williamson in \([\text{Wil03}, \text{page 415}]\) stresses that Quine’s conception of ontology also requires absolutely general, unrestricted quantification\(^{24}\).

However, I think that the defence of absolute generality provided by such recent thinkers is often weaker than that of Cartwright, as they generally concede too much to the opponents of absolute generality. Many of them think that the possibility of absolute generality has some very strong presuppositions; they think that it requires that there must be a kind of higher-order quantification that is not reducible to first-order quantification or a revision of standard propositional logic. I want to argue, however, that the possibility of absolute generality does not require either of these assumptions (though they might, of course, still be true for other reasons). Indeed,

\(^{23}\)I do not mean that using unrestricted quantification would be sufficient for a statement being absolutely general, only that it is necessary. A statement can only be absolutely general if the predicates that occur in it must also be very general, indeed categories.

\(^{24}\)Chalmers also holds in \([\text{Cha09}, \S 9, \text{page 105}]\) that ontological statements cannot be restricted to a domain, but holds that they must be restricted to a world; Chalmers thinks that they demand that there is an absolute domain, the domain of the actual world. I will argue, however, that ontological statements do not require the existence of any absolute domain either.
one of the most prominent proponents of absolute generality, Quine (whose commitment to it Cartwright shows) thought that higher-order quantification that was not reducible to first-order quantification was not possible. Quine sometimes went so far as to deny the intelligibility of higher-order quantification even if it was held to be reducible to first-order quantification, and I think that this was highly implausible; however, I think that there is a good case against a kind of higher-order quantification that would not be reducible to any kind of first-order quantification (and not just to standard first-order quantification).

The kind of quantification that Tarski used in his theory of truth was unrestricted quantification. Tarski thought that variables could range over all classes. This can be confirmed if we look at Tarski’s treatment of quantification in [Tar83a, page 190], where Tarski is being most precise; here Tarski says for example that

\[ \text{for all } a, a \text{ satisfies the sentential function } \bigcap_2 t_{1,2} \text{ if and only if for all classes } b \text{ we have } a \subseteq b. \]

Note that Tarski speaks here unrestrictedly of \textit{all} classes, not just of all classes belonging to the domain of a model, as would be typically done in modern model-theory.\footnote{There is one way in which the quantification used by Tarski and that used by Williamson, Weir and many others is not wholly unrestricted. Tarski only quantifies over classes in the examples he gives. This is not because he would think that only classes can be quantified over; it is only one example he uses out of many he could have chosen. Tarski does think that quantification must be restricted typically; he is still thinking in terms of Whitehead’s and Russell’s type theory. However, this is an entirely different and far weaker restriction than the restriction to members of the domain of a model. Hilbert apparently shared this restriction but thought that there had to be a further restriction within every type; by relativising quantification over individuals to some individual domain all class types are at once also relativised to classes that are classes of those individuals or classes of those classes of those individuals or \ldots Cartwright and Williamson quantify only over individuals; the typical restriction they use is the reverse of that used by Tarski. I will argue that quantification can be unrestricted also in the stronger sense that it need not be restricted even typically; we can quantify at once over all entities of all types.}

A whole book has been gathered containing articles on this topic, namely [RU06], edited by Agustín Rayo and Gabriel Uzquiano. The possibility of metaphysics and ontology in the traditional sense depends on the possibility of unrestricted quantification i. e. of absolute generality, so it is one of the most important metantological or metametaphysical questions and I must deal with this topic.
The best justification of the claim that we can quantify unrestrictedly i.e. over absolutely everything, is that denying it appears to be self-defeating; you have to presuppose the truth of the claim in order to deny it. In saying that we cannot quantify over absolutely everything, a person is himself quantifying over absolutely everything. If he is not, he is just denying that we can quantify over some limited domain of entities, and this is a far weaker claim.

Cartwright stresses that the claim that we can quantify over everything does not imply that there would be a universal domain over which we can quantify which would be a set or a class. A universal domain would have to contain all classes (as well as whatever Urelements there might be) and there are reasons to suspect that this would lead to paradoxes and hence be contradictory. While standard set theories such as ZF do not allow there to be a universal class, there yet are set theories such as Quine’s NF (presented in [Qui53c]) and a variant of it called NFU which allow there to be a universal set, and I am quite attracted to them; nevertheless, the possibility of unrestricted quantification does not require the truth of such theories. If it did this would obviously put that possibility in a perilous position; I want to argue that the possibility of unrestricted quantification is independent of what set theories are correct, since it is a highly general logical or semantical claim, more general than set theories.

The claim that absolute generality would imply the existence of a universal set is based on a principle that to quantify over certain objects is to presuppose that those objects constitute a collection or a completed collection. Cartwright calls this principle in [Car94, §IV, page 7] the All-in-One Principle. Cartwright thinks that the principle is false; of course, opponents of unrestricted quantification often think that it is correct. Some adherents of unrestricted quantification such as Alan Weir seem to accept this principle, but this makes the defence of their position needlessly hard.

Alan Weir asks in [Wei06, page 336] whether we do not need domains in order to achieve the goals for which the formalization of language is a means. According to him one goal is to give a systematic account of the meaning of important logical notions such as 'every'. He says:

We semi-formalize some mathematical English into a sentence such as 'for every number $x$ there is some number $y$ with $x < y$'. Our account of truth for generalizations entails that this
universal generalization is true relative to an assignment \( \sigma \) to variables just in case for every \( x \)-variant assignment \( \sigma(x/\alpha) \) like \( \sigma \) except, at most, that it assigns \( \alpha \) from the domain \( D \) to \( x \), the open sentence there is some number \( y \) such that \( x < y \) is true relative to \( \sigma(x/\alpha) \). How can we give such an account unless the definition of truth generalises to truth relative to all variant assignments? Can we do this without domains?

Weir thinks that we can, but only by using higher-order quantification. He makes use of an idea of George Boolos, which replaces the notion of an interpretation function with the notion of an interpretative relation, where interpretative relations are quantified over in the metalanguage with predicate variables such as \( R \). I think that we can do it even without making use of higher-order quantification. Of course I do not think that there is anything wrong with making use of predicate variables in a semi-formalized metalanguage or the notion of an interpretative function; these elaborate manoeuvres just are not necessary. Ultimately, Weir thinks that even this higher-order approach is not enough, and thinks that we have to revise logic in accordance with Kripke’s theory of truth. I find it quite plausible that some such revision of logic might in the end be necessary in order to handle the paradoxes; however, I do not think that it is necessary in order to defend the possibility of unrestricted quantification.

In fact giving a systematic account of the meaning of logical constants like ”every” without using domains is ridiculously simple. The definition of truth does generalise to truth relative to all variant assignments, but this does not require the use of domains or the use of higher-order logic in the metalanguage as in Weir’s theory. All we have to do is remove the reference to domains from the truth-conditions that Weir gives. Any function from the variables of the language to objects can be taken to be an assignment, not just those whose values belong to some domain. This is what I will do in most of the remaining dissertation when I give formal semantical rules. We can then say in the case of the example that Weir gives that this universal generalization is true relative to an assignment \( \sigma \) to variables just in case for every \( x \)-variant assignment \( \sigma(x/\alpha) \) like \( \sigma \) except, at most, that it assigns some object \( \alpha \) to \( x \), the open sentence there is some number \( y \) such that \( x < y \) is true relative to \( \sigma(x/\alpha) \). Indeed, this is what Williamson does in [Wil03, page 418], so it is strange that Weir does not see that this is the
Williamson has, however, a reason why he thinks that this kind of solution is not sufficient; Linnebo has a similar argument in [Lin06, §6.2, page 150] which he calls the semantic argument. Williamson argues in [Wil03, page 425] that sooner or later, the theorist will want to generalise over all interpretations of various forms in the language. He goes on to argue in [Wil03, page 426] that when we apply the definition of logical consequence, it must be possible to interpret a predicate letter according to any contentful predicate. Thus, whatever contentful predicate we substitute for 'F', some legitimate interpretation (say, $I(F)$), interprets the predicate letter accordingly. Williamson then derives a contradiction from this assumption by defining a verb 'R' so that for everything $o$, $o Rs$ if and only if $o$ is not an interpretation under which $P$ applies to $o$ and then putting 'R' in place of 'F'. We get from this that for everything $o$, $I(R)$ is an interpretation under which $P$ applies to $o$ if and only if $I(R)$ is not an interpretation under which $P$ applies to $o$, and therefore $I(R)$ is an interpretation under which $I(R)$ applies to $o$ if and only if $I(R)$ is not an interpretation under which $I(R)$ applies to $o$.

Williamson says that this paradox is a variant of Russell’s paradox, but it employs no notion of set, class or domain. However, it seems to me that the paradox at which Williamson arrives is instead a variant version of the old Liar paradox, or perhaps a combination of the two paradoxes. Williamson’s assumption that it must be possible to interpret a predicate letter according to any contentful predicate of the meta-language violates Tarski’s discovery (expressed for example in [Tar44, pages 351,352]) that the metalanguage must be essentially richer than the object language. Tarski tells us that if the

---

26In any case it is far from clear that Williamson’s paradox does not employ a notion of class, since if interpretations are taken to be objects they are usually taken to be classes; interpretations are generally understood as classes of pairs consisting of expressions of the object language and other entities. We can doubt whether any coherent and systematic understanding of the notion of interpretation is possible that does not make it to be a class or an intensional relation which has a class as its extension. We can argue that in the case of many contentful predicates 'F’ of the metalanguage, there is no class of all entities that are F (or even an abundant property that would be the intension of the predicate) and in this case there can be no pair of which this class would be a member, and therefore there can be no interpretation that would contain such a pair as one of its members. This is in fact obviously true of such predicates as 'self-identical' in a standard set theory such as ZF and can well be true of the verb R that Williamson and Linnebo define. This is enough to avoid the paradox that Williamson arrives at while continuing to maintain that we can quantify over absolutely everything.
condition of essential richness is not satisfied, it can usually be shown that an interpretation of the object language in the meta-language is possible. However, this implies that it is possible to reconstruct in the metalanguage the paradox of the liar. This seems to be exactly what Williamson does. Indeed, we can derive a paradox from Williamson’s assumption even without assuming that interpretations are objects or even using the concept of interpretation at all; just let ‘F’ be the truth predicate! We can indeed associate a predicate with every entity of which we can speak in the metalanguage, as we must do if quantification is to be genuinely unrestricted, but not with every predicate of the metalanguage, for the metalanguage cannot speak of all of its own predicates, if it is not to be semantically closed. However, this restriction has nothing to do with a restriction to domains; quantification in the object language can be absolutely unrestricted (as we have seen it was in Tarski’s theory) even if the expressive power of the object language is restricted (as it was in Tarski’s theory). Therefore this difficulty has nothing to do with unrestricted quantification, so the further complications which Williamson introduces to his theory are quite unnecessary.

Many arguments have been given against the possibility of unrestricted, absolutely general quantification. Rayo and Uzquiano distinguish five primary arguments; however, it seems to me that when we consider them carefully, they can be reduced to three. None of them seems at all strong to me; one of them would prove too much and would lead to a blatant absurdity, and the premises of the remaining two are no more plausible than their conclusion. There is no space here to go through all the arguments carefully; I will only sketch where their weaknesses seem to me to lie.

One argument is according to Rayo and Uzquiano in [RU06, §1.2.1, page 4] supposed to be based on the thought that certain concepts such as ‘set’ or ‘ordinal’ are indefinitely extensible. However, the indefinite extensibility of concepts is surely compatible with absolute generality, for absolute generality is about the interpretation of quantifiers, while indefinite extensibility is connected with the interpretation of predicates, whose semantic values concepts are. In fact we need some principle such as the All-in-One principle in order to get from indefinite extensibility to an argument against the possibility of unrestricted quantification: Rayo and Uzquiano call the argument based on the All-in-One Principle in [RU06, §1.2.2, page 6] a related argument, but it is really basically the same argument as the argument
from indefinite extensibility. However, Richard Cartwright has already in my view shown that we have no reason to accept that principle, so we can eliminate the first two of the arguments Rayo and Uzquiano distinguish.

A third argument is supposed (see [RU06, §1.2.3, page 9]) to be an argument from reconceptualization. This argument is supposed to be based on ontological relativism of the kind presented by Carnap in [Car50]. I will present arguments against this kind of relativism in Section 5.9 of this work, to which I here refer the reader.

The fourth argument is supposed to be based on semantic indeterminacy. Both the third and the fourth argument are based on the same basic relativistic ideas derived from Carnap, Putnam and Quine, so I rather doubt if they can be considered genuinely separate arguments. However, the conclusion of this argument is not so much that all quantification is restricted but rather that there is no fact of the matter about whether quantification is restricted or unrestricted. The premises of the argument prove not only that we cannot quantify over absolutely everything, but rather that we cannot quantify determinately over anything at all, i.e. that we cannot quantify at all. The extreme radicalism of the conclusion of this argument is already by itself a reason to reject this argument. The argument is self-defeating.

---

27 None of the arguments Rayo and Uzquiano present is said by them to be based directly on verificationism. Yet the third and fourth arguments have a verificationist background (as well as a structuralist background) and hence represent a remnant of logical positivism. This is shown by the fact that Rayo and Uzquiano refer to Hilary Putnam’s article [Put80] as one of the writings in which the fourth argument is set out, and in this article Putnam explicitly says (on [Put80, page 464]) that he opts for verificationism. Many absolutely general statements are indeed very likely to be unverifiable according to any strong kind of verificationism; we have seen that even relatively general statement are unverifiable according to very strong kinds of verificationism. It does appear impossible that the understanding of the truth conditions of absolutely general quantification could ever be manifested (in the strong sense implying the possibility of conclusive verification of claims that such quantification is understood). However, we have seen that there are good arguments against verificationism, so this is no good reason to reject absolute generality. As we have seen, there are reasons to think that even the understanding of restricted quantification as standardly conceived could not be manifested, and indeed these reasons appear to indicate that the understanding of any kind of objectual quantification could be manifested (in the strong sense) at all.

28 Indeed, one of the articles to whom Ray and Uzquiano refer, Quine’s [Qui68], clearly implies that semantic indeterminacy is not incompatible with absolute generality. Quine argued in [Qui68, pages 202-203] that we had better beware of repudiating universal predication. Universal predication, however, implies absolutely general quantification, as universally predicing a predicate P involves quantifying over absolutely everything and then predicating P of these entities quantified over.
and could not even formulate the argument itself.

The fifth is an argument from sortal quantification; it says that all quantification must be sortally restricted, as there must be a sortal term that gives an identity criterion for the entities quantified over. This argument is better than the previous one since it is not obviously self-defeating; however, it seems again to be question-begging. Michael Jubien argues rather convincingly in [Jub96] that the demand for identity criteria is ill-conceived with respect to any entities at all. Obviously if he is correct the argument from sortal quantification collapses rather dramatically. However, even if we admit that all entities must have a criterion of identity, we can hold that we can quantify over entities that have a criterion of identity without explicitly mentioning this criterion of identity. The idea that quantification can be unrestricted or absolute is linked, as e.g. Timothy Williamson has argued in [Wil06], to the idea that identity can be absolute. Many philosophers who think all entities must have an identity criterion yet agree that identity can be absolute and not just relative to sorts. David Wiggins, for example, has held this kind of position in [Wig01, pages 22-24]; Wiggins distinguishes from each other two theses, Sortal Dependency of Individuation and Relativity of Identity, and argues that the first is correct but the second is false. However, only the second, Relativity of Identity, would be incompatible with absolutely general quantification. Therefore if either of Jubien’s or Wiggins’s very different views are correct then the fourth argument fails. I cannot here go deeper into the semantics of sortals or the problem of identity criteria, but this suffices to show that there are principled reasons to doubt the validity of the fourth argument.

5.7 Explicating Quine’s criterion: Different Kinds of Ontological Commitment

We have seen that Quine’s suggestion can be motivated historically. However, this does not suffice to dissipate all problems relating to it, since the suggestion must be clarified. Quine’s formulations suffer from at least three different kinds of obscurity or non-equivalence: it is obscure whether the criterion concerns explicit or implicit commitment, it is obscure what kinds of entities carry commitments and finally, most seriously, it is obscure whether the whole criterion is a syntactic or a semantic criterion. Indeed, Quine’s
formulations do not even make it fully clear whether his theory is a metaontological theory or a first-level ontological theory, since some of his formulations suggest that his theory would be a theory of being or existence rather than of ontological commitment.

As Charles S. Chihara has shown in [Chi74, page 71], Quine says sometimes that his criterion is a criterion for determining what a theory presupposes there is, at other times that it is a criterion for determining what a theory implies there is and at others that it is a criterion for determining what a theory says there is. As Chihara says, these are in general different things. We can say that an assertion or sentence is explicitly committed to some entities if it says that there are such entities, but a theory or an assertion is implicitly committed to some entities if it implies or presupposes that there are such entities. I will be interpreting the potentially ambiguous statement that an assertion says something very strongly, so that if an assertion says something then its explicit ontological commitments will be immediately clear to anyone who can understand the statement; he will not even have to perform any deductive reasoning to find them out. I will not be making a sharp distinction between the two kinds of implicit presupposition, the case where an assertion implies that there are some entities and the case where an assertion presupposes that there are some entities. The reason for this is that while the distinction between implication and presupposition is important for natural languages, it does not exist in most artificial languages, and the discussion regarding ontological commitment usually takes place with respect to artificial languages such as the language of predicate logic (and I think that it should take place primarily with regard to them); however, I will in Section 5.15 address the important and difficult question of the ontological commitments of natural language assertions, and I will then look at this distinction. I will suggest that we can find in Quine both a

\[29\] Howard Peacock has also stressed the importance of a distinction between explicit and implicit ontological commitment in [Pea11]. However, Peacock does not understand quite the same distinction with these words than I do. Peacock says in [Pea11, page 80] that whether a fragment of language bears explicit ontological commitment depends on the beliefs and intentions with which it is used. Peacock himself notes in [Pea11, page 85] that the property of being explicitly committing (in his sense of the words) is not a semantic property. Peacock therefore uses the words “explicit ontological commitment” to refer to what I will call pragmatic ontological commitment. However, I intend to use both words to stand for varieties of semantic ontological commitment, so that neither the explicit nor the implicit commitment carried by a sentence depends on the beliefs and intentions with which it is used but only on the semantic rules of the language.
criterion for explicit commitment and a criterion for implicit commitment; however, these cannot be the same criterion. They must be distinguished from each other more carefully than Quine himself did. I will give definitions for both explicit and implicit ontological commitment. I will divide implicit ontological commitment further into formally implicit and materially implicit ontological commitment based on whether the implication or presupposition in question is purely formal, logical implication or some weaker kind of implication, material or analytical implication. I will further distinguish different kinds of materially implicit ontological commitment.

Quine’s formulations are unclear even with regard to what it is that he is trying to explicate. Quine is not quite consistent regarding what kind of entities have ontological commitments. In different places he gives different suggestions about what entities carry commitments. Quine often speaks about people being committed to something, and this suggests what can be called ontological commitment in a pragmatic sense. However, according to [Qui53b, page 103] the criterion applies primarily to discourse rather than men; according to a famous and relatively clear formulation,

entities of a given sort are assumed by a theory if and only if some of them must be counted among the values of its variables in order that the statements affirmed in the theory be true.

Sometimes semantical and pragmatical ontological commitment are distinguished by the use of different verbal forms; philosophers often speak of people being committed to some entities but statements carrying commitment to some entities. I will follow this practice most of the time; however, it must be recognized that it is not the only legitimate way of using these expressions. Quine and philosophers purporting to follow him also apply the criterion to individual sentences or assertions besides full theories. It is surely not impossible that all of these entities would have ontological commitments in some sense, but it seems unlikely that so different entities could have ontological commitments in exactly the same sense. Also such words as ”theory” or ”assertion” can be used in different senses. Quine, however, does not investigate the relationship between these concepts of ontological commitment at any length, and we can argue that because of this he does not avoid confusing them with each other\(^\text{30}\).

\(^\text{30}\)There are also other suggested carriers of ontological commitment that Quine does not
Theories are usually understood\textsuperscript{31} to be logically closed sets\textsuperscript{32} of statements\textsuperscript{33}. Not every such set, however, is plausibly called a theory; a theory is more plausibly the logical closure of a set of statements all of which have been asserted or believed in by some person or group of persons\textsuperscript{34}. This even consider, because his philosophy as a whole is hostile to them, such as propositions; I will later touch upon this possibility in a later note.

\textsuperscript{31}E. g. already in [Car48, §9, page 34] Carnap says that a theory may be regarded as the class (in general transfinite) of all those sentences which are deducible from a given finite set of sentences, e. g. physical laws.

\textsuperscript{32}In saying this I am not committing myself to the Received View; I suppose that a theory will have to be totally interpreted, not just partially, and I am not supposing that it would have to be axiomatizable in extensional first-order logic, but allow it to be formulated in higher-order and intensional languages. Nor is it necessary to even think that a theory would have to be axiomatizable at all; the demand of logical closure could be formulated in fully semantical terms, so that if $T$ is a theory and $\phi \in T$, and $\phi \models \psi$, then $\psi \in T$. A theory could also be taken to be a set of propositions; however, this would make it easier to formulate the theory of ontological commitment, and I will here try to carry through the harder task of formulating the theory so that even those who do not believe in the existence of propositions could accept it.

\textsuperscript{33}The demand of logical closure is probably in the end too strong; however, it is a useful or perhaps even necessary idealisation when we begin to develop the theory of ontological commitment. After all, scientific theories are commonly understood as logically closed in the philosophy of science and the ontological commitments of scientific theories are surely among the most important of ontological commitments, even if they are not the only important ones as Quine himself thought. However, the assumption of the closure of scientific theories notoriously causes many problems in the philosophy of science (even when explicitly ontological questions are not addressed), and for similar reasons it causes problems also for the theory of ontological commitment. Just as an inconsistent sentence according to classical logic logically implies all sentences, so the assumption of logical closure leads us to suppose that an inconsistent theory carries implicit ontological commitment to all (possible and even impossible) kinds of entities. However, there surely appears to be a sense in which even an inconsistent statement or theory carries implicit ontological commitment to some entities but not to all entities. Nevertheless, it has proved difficult to find any viable alternative to the view of theories as logically closed. It is of course possible to take theories to be just arbitrary (and naturally finite) sets of statements without any closure at all, and my theory of ontological commitments can easily handle such a view of theories, since in this case the set of the ontological commitments of a theory could just be identified with the set of the explicit ontological commitments of all its sentences. However, this view probably would individuate theories too finely, confusing a formulation of a theory with that theory itself. It might be best to use some kind of paraconsistent or relevant logic, and it is surely possible to define a weaker concept of implicit ontological commitments than the one I define in this dissertation by using a paraconsistent or relevant notion of logical implication instead of classical logical implication. However, I cannot go any deeper into this problem in this dissertation.

\textsuperscript{34}Often it is also demanded that it be the closure of a big set of statements; paradigmatic examples of theories are such big theories as the Newtonian theory or the Theory of Relativity. There are many different theories about how big a theory would have to be, but none of them seem plausible to me. Strong semantic holists would say that the minimal unit of meaning is the entire set of all statements believed in by a person or
suggests a connection between statements (or sentences) and theories; to every statement and every sentence there corresponds a miniature theory, namely its logical closure. This can be used to yield a connection between the ontological commitments of theories and statements; a statement carries (implicit) commitment to some entities if and only if the theory which is its logical closure carries ontological commitment to them and a theory carries ontological commitment to some entities if and only if some statement contained in it carries (explicit) ontological commitment to them.

Ontological commitment in a pragmatic sense differs more from these two notions than they do from each other. However, pragmatic ontological commitment can also be understood in different senses. In one sense a person can be said to be ontologically committed to some entities if some statement he makes seriously or some theory he supports carries commitment to them when it is interpreted in the standard way. In another sense a person can be said to be ontologically committed to some entities if some statement which he makes would carry commitment to them if it were interpreted as he intends to interpret it. These notions of pragmatic ontological commitment can diverge when a person misunderstands the language he is using. In a third sense we can also speak of pragmatic ontological commitment carried by conversational implicatures.

It must be noted that it follows from this definition that a person need not always be ontologically committed in the pragmatic sense to entities to which a statement he utters carries ontological commitment semantically, if he does not make that utterance seriously (as is also stressed by Michaelis Michael in [Mic08, pages 57-59]. For example, if the person is acting in a group of persons. However, since such strong semantic holism is very implausible (as I have already argued) we can ignore this demand (and indeed we must, since such a holism would make it impossible to develop any semantic theory, whether of ontological commitments or anything else). Quine himself in his latest period would have thought that only a set of sentences that implies observation categoricals is a theory; however, this was due to his verificationism, which I have argued we should abandon, so this does not seem to me to be good reason to restrict theories in this way. Sometimes theories are only understood as sets of statements concerning unobservable entities, as in the common contrast between observational and theoretical; however, this restricted use of the word is not suitable for my purposes here. Observation sentences themselves clearly carry ontological commitments to observable entities, whether physical entities, mental states or sense-data. Another common restriction of the word "theory" is to sets of statements containing synthetic statements; however, we surely often speak about mathematical theories, which according to many philosophical accounts consist solely of analytical statements. I will not be using the word in this restricted sense either.
stage play or movie, then it is not he or she, but the fictional character he plays, who is committed to those entities. An example that is important to the philosophy of science is that someone who interprets a scientific theory instrumentally need not be ontologically committed to the entities to which the theory carries ontological commitment; for instance, van Fraassen is not committed to the entities to which physical theories carry ontological commitments, though he in some weak sense accepts those theories. As we have seen, Quine himself gave an instrumentalistic interpretation to scientific theories in [Qui53e], so even though he thought that scientific theories carried ontological commitment to physical objects and classes, he was not at that stage of his philosophical development committed to these entities, though there are plenty of indications that he later rejected this instrumentalism and became so committed.

This may not be a complete division of the senses of pragmatic ontological commitment or fully adequate even so far as it goes; however, I cannot here go deeper into the pragmatics of ontological commitment.

Quine's famous slogan - to be is to be a value of a bound variable - might suggest that Quine's theory would be a definition of what being itself is. However, it is clear that if Quine's slogan were interpreted in this way, it would be circular; according to the theory itself the indefinite article "a" already expresses being, so the notion of being would occur in the definiens as well as the definiendum. It is rather clear that Quine did not intend his slogan to be a definition, analysis or explication; he says in [Qui69, page 97] that it is unreasonable to ask for an explication of existence in simpler terms. Besides, the slogan suggests that the notion of being would be a semantical notion, which is clearly incorrect, so the slogan is defective even if not taken as a definition.

In fact Quine's theory implies a kind of theory of being or existence, but it is better expressed slightly differently than the slogan would suggest. I will suggest a new slogan that may be less misleading. Rather than saying that to be is to be a value of a bound variable, we should say that to be is to be something. This saying can be motivated by Quine's own words; e. g. he says in [Qui69, page 94] that "a is" is short for "a is something." More specifically, to exist is to be identical with something. However, this slogan, like all slogans, is yet not free from ambiguity (as the notion of "something" can be understood either objectually or substitutionally or in other ways),
and is no substitute for a technical theory.

The notion of something, the notion of particular quantification, must itself be taken as a primitive; any attempt to define it would lead to circularity. In a way then being itself must be taken as a primitive, since the theory holds it is the same as the notion of particular quantification.

5.8 Justification of Ontological Commitments and Pragmatism

As Quine himself emphasised in [Qui53d, page 16], his criterion of ontological commitment is meant for the recognising of the ontological commitment of discourse, and is separable from his views of how to adjudicate between rival ontologies, i.e. of how ontological commitment is to be justified.

Quine claims in [Qui53d, page 16] that we adopt, so far as we are reasonable, the simplest conceptual scheme into which the disordered fragments of raw experience can be fitted and arranged.

The presupposition here that our experience is disordered and fragmentary is phenomenologically quite dubious, and suggestive of an unattractive constructivist view according to which we impose upon the world what order there is in it, which fits badly with the scientific realism Quine adhered to in other writings. Surely there is at least some order already given to us in our experience and not just imposed upon it. Nevertheless, Quine’s basic idea is quite plausible, if taken in the form that we adopt the simplest theory in which our experience can be further ordered, though even in that form far from uncontroversial.

This claim of separability is important, as it clearly separates semantic and epistemological questions relating to ontological commitment, and so escapes the confusion between semantic and epistemological questions that was characteristic of logical positivism. However, it does not really fit well together with Quine’s semantic holism, which reflects the part of Quine’s philosophy still indebted to logical positivism. Semantic holism implies that we might change our views of the ontological commitments of discourse in order to preserve our views of how to adjudicate between rival ontologies, but by separating these questions Quine here rejects this implication of his semantic holism.

Jody Azzouni distinguishes in several writings between criteria for what a discourse is committed to and criteria for what exists, devoting separate chapters for these two themes in [Azz04]. This seems very similar to Quine’s distinction; however, may not be quite the same. Quine’s criteria for adjudicating between rival ontologies are epistemic, but as Azzouni stresses, his criteria for what exists are purely metaphysical. There are then three kinds of criteria that might be distinguished.

I have argued that the Eleatic principle is best understood as a principle how to
Quine’s view of how to recognize the ontological commitments of discourse, while rejecting his view of how to adjudicate between rival ontologies. Quine himself recognizes that there are different types of simplicity which can conflict - e.g. the scientific and especially mathematical theories to which Quine himself is committed are in some ways far from simple, as they postulate lots of unobservable entities - so even if his answer is acceptable as far as it goes it is far from a complete answer. I would say rather that we adopt the ontology which is most likely to be true on the basis of our experience. I will come back to this question in Section 6.1 of this work.

In later articles Quine held the view that decision between rival ontologies is to be done on pragmatic criteria. E.g. in [Qui53a, page 79] he said that our standard for appraising basic changes of conceptual scheme must be a pragmatic standard. This is often taken to be an essential part of his theory of ontological commitment. However, there is no mention of pragmatism in [Qui53d], and this view can be taken as an afterthought of Quine’s which can be separated from the main body of Quine’s theory of ontological commitment, from both of the two parts of his theory I distinguished above. At least we have been given no reason to think that the use of simplicity in adopting our theory would be necessarily connected to the use of a pragmatic standard. The simplest theory consistent with our experience can yet contain many statements which are of no practical importance.

Though Quine refers to pragmatism approvingly in many places, it is yet not sure that he really was a pragmatist. It has been argued (e.g. by Heikki J. Koskinen and Sami Pihlström in [KP06]) that Quine himself was not clear what it means to be a pragmatist, and that he cannot be classified as a pragmatist in any strong historical sense. Quine’s pragmatism, as evidenced in passages such as the one I have referred to, differs in an important way from traditional pragmatism. Quine presents it as an epistemological or methodological principle, not a semantic one, as James’s pragmatist theory of truth (as seen in [Jam48]) and Peirce’s pragmatist criterion of meaning (as presented in [Pei78]) were. Quine’s own disquotational theory of truth differs from a pragmatic theory of truth. This makes Quine’s pragmatism more defensible than the traditional pragmatic theories (or than the verifica-
tionism I have argued Quine himself advocated in [Qui53c]). Nevertheless, I
would not be willing to subscribe to it. One big problem with Quine’s prag-
matism is that if taken as a criterion for the justification of all ontological
commitments it carries the risk of circularity; in order to know the accep-
tance of what theories is useful for us, we must know the truth of at least
some theories, namely theories concerning experience, and in this case we
must already be able to justify some ontological commitments, namely com-
mitments to observable entities, whatever they are. At most a pragmatic
criterion could justify our commitments to theoretical entities.

It is sometimes suggested that Quine’s espousal of pragmatist ideas some-
how makes his view anti-metaphysical. Huw Price has used this kind of
argument in many articles; e. g. he says in [Pri09, page 327]:

Quine’s criticism of Carnap cannot provide vindication of tra-
ditional metaphysics, for if all issues are ultimately pragmatic,
there can’t be the more-than-pragmatic issue of the kind the
metaphysician requires.

However, I will argue that there is nothing anti-metaphysical in pragmatism
as such, if the word ”pragmatic” is understood at all in the sense it is
ordinarily used in philosophy. Therefore even if Quine was a pragmatist,
this would not prevent him from vindicating metaphysics.

Peirce’s pragmatic criterion of meaning as presented in [Pei78] is indeed
highly similar to the verificationist criterion; such formulations as

Our idea of anything is our idea of its sensible effects

in (CP5.401) (collected in [Pei92, page 132]) make this very clear (though it
is not clear whether strong formulations such as this are consistent with all
of Peirce’s other opinions or whether Peirce may not have rejected such a
strong criterion at other times of his philosophical development). However,
the pragmatic criterion was generally used by pragmatists not to condemn
all metaphysics as such, as logical positivists used the verificationist crite-

ri-on, but rather to argue for a certain variety of metaphysics, as such his-
torians of verificationism as Misak recognize (see [Mis95, pages 118-119]).
I have already argued that even the verificationist criterion of logical pos-
itivists rather implied a certain kind of metaphysics than condemned all
metaphysics, as indeed such later philosophers as Bergmann and Dummett
recognized; however, in the case of the pragmatists there is no need to argue this, as they themselves recognized this from the start, unlike the logical positivists.

Several classical pragmatists like Peirce and James engaged in extensive metaphysical speculation and were if anything willing to indulge in what most analytical metaphysicians would consider too far-fetched metaphysical speculations, such as Peirce’s evolutionary metaphysics (with doctrines such as tychism, synechism and agapism) in writings such as a series of articles on metaphysics and cosmogony in the journal Monist, including [Pei83] or the panpsychist cosmology at least considered seriously in James’s [Jam09]. The attitude of classical pragmatists toward metaphysics may indeed have vacillated between more and less favourable, but it is by no means clear that their more hostile statements would be more faithful to the true essence of pragmatism than the more friendly ones. Many modern pragmatists are also committed to taking metaphysics seriously, as seen for example in Sami Pihlström’s pragmatist metaphysics in [Pih09]. Pihlström argues at length for the compatibility of pragmatism with metaphysics in [Pih09] as well as together with Koskinen in [KP06]. William T. Myers has also argued for the compatibility of metaphysics and pragmatism in [Mye04], opposing Charlene Haddock Seigfried, who claimed in [Sei01] that pragmatists should get rid of the word "metaphysics". There have indeed been pragmatists who have been hostile to metaphysics, but this is clearly not something that would

---

38 Peirce often defended or presupposed panpsychism, an idealistic theory, e.g. reminding us in (6.301) (see [Pei83, page 187] and [Pei92, page 361]) that all matter is really mind, as if this was well known to the reader. James also initially inclined towards idealism; he criticized in the second and third lecture of [Jam09] the kind of monistic idealism that had ruled Anglo-Saxon philosophy and that Russell and Moore had also criticized, but instead of replacing it with realism he instead argued in favour of a pluralistic idealism. He discussed at length in the fourth lecture of [Jam09] a panpsychist theory propounded by the German metaphysician Fechner and considered it (if made slightly more pluralistic still) at least probable. However, James’s [Jam22] instead followed the early logical positivists in defending neutral monism as the solution to the problem of the relation between mind and matter. James talked of a world of pure experience, using the phrase coined by Richard Avenarius. I have already argued that neutral monism is a metaphysical theory, so both of James’s theories count as definitely metaphysical. William T. Myers compares James’s metaphysics to Whitehead’s process metaphysics in [Mye04]; this holds more of James’s earlier than his latter theory.

39 John Dewey was more wary of metaphysics than the other three major classical pragmatists, but even he at times presented a metaphysics of experience, e.g. in [Dew29]. Richard Rorty’s views are often given as an example of pragmatism’s hostility to metaphysics. However, many philosophers have challenged Rorty’s claim to represent the pragmatist tradition. E.g. Susan Haack (who is herself in the pragmatist tradition) de-
follow from their pragmatism as such but from principles specific to their peculiar form of pragmatism. There is therefore no reason to think that a metaphysician as such would require a more than pragmatic vindication of any (metaphysical or physical) theory. Therefore even if pragmatism were considered an essential part of Quine’s theory of ontological commitment, this would do nothing to prevent Quine’s theory from providing vindication of metaphysics, contrary to what Price without any justification presupposes.

However, while pragmatism has never been opposed to metaphysics as such, pragmatism has indeed been generally associated with the rejection of a realistic metaphysics and the acceptance of a non-realist metaphysics. Such modern pragmatist metaphysicians as Pihlström explicitly oppose realism (or at least strong forms of realism such as metaphysical realism, but it can be argued that weaker forms of realism such as Putnam’s internal realism are not genuinely realistic). There have indeed been philosophers who have tried to reconcile pragmatism with realism - Peirce’s version of

---

\[^{40}\text{Nevertheless, Quine’s criticism of Carnap may not suffice for the vindication of metaphysics; I will argue later in Subsection 5.9.2 of this dissertation that it will have to be supplemented with additional considerations directed against Carnap’s relativism. Also the differences of Carnap and Quine in their attitude to ontology did not depend essentially on their views about analyticity, as Price supposes, but rather on their different theories of quantification, as I will show in the same section.}\]

\[^{41}\text{One complication is that in one passage to which Price refers Quine is not formulating a pragmatic idea, as Price claims, but an instrumentalist one (concerning the justification of claims about physical objects and classes). Instrumentalism is not the same as pragmatism, but rather incompatible with at least the most famous forms of pragmatism. An instrumentalist about a theory says that the theory is useful but not literally true. However, a pragmatist - in the common sense of one that accepts the pragmatic theory of truth - would say that there is no distinction between useful and literally true theories, since according to that theory truth just is usefulness. Price seems then to confuse instrumentalism with pragmatism, and it is possible that the same confusion already existed in Quine, unless Quine vacillated between the two views. However, it would make no sense to be an instrumentalist about all possible theories, for if it makes sense to call some theories mere instruments, other possible theories must be more than instruments. Even in his instrumentalist mood Quine implicitly presupposed an ontology of sense-data as literally true, so that even then he presupposed a metaphysical theory.}\]

\[^{42}\text{Therefore Price’s claim might be justified if he understood ”traditional” metaphysics so narrowly that it only covered such old realistic metaphysics as the basically realist Aristotelian tradition. However, that would be an unnaturally narrow use of the word; surely such non-realistic metaphysics as post-Kantian German idealism (which already usually had a pragmatic character as it commonly relied on the Kantian concept of practical reason while denying Kant’s strict division between theoretical and practical reason) would generally be taken as instances of traditional metaphysics.}\]
pragmatism, pragmaticism, especially has often been seen as close to scientific realism - but it is controversial how far such an attempt can succeed. Such statements of Peirce as the above quoted formulation of the pragmatist criterion of meaning obviously pose a difficulty for such a project, though it is possible to argue that Peirce’s views changed and some of them are more reconcilable with scientific realism than others. Quine’s appeal to pragmatism is indeed weaker than the traditional pragmatism, as I have argued above, being epistemological rather than semantic, so it might be compatible with realism even if traditional pragmatism is not. However, if this is not the case, a realistic metaphysician may indeed want to reject this part of Quine’s theory. Nor is there any reason why he could not do this while retaining Quine’s theory for the ontological commitments of discourse.

5.9 Objectual and Non-objectual Quantification

The last obscurity in Quine’s formulations is most crucial. When Quine says that the use of bound variables is the criterion of ontological commitment, he seems to suggest that the criterion is syntactic, since the notion of a bound variable is most naturally understood as syntactic. Quine’s criterion of ontological commitment is often called a syntactic criterion and e. g. Jody Azzouni in [Azz04, page 50] argued for a syntactic criterion. However, the notion of ontological commitment is clearly on the face of it semantic (if not pragmatical), and there are obvious reasons to doubt whether it can be reduced to syntactic notions. There are also reasons to doubt whether Quine himself seriously meant for the criterion to be purely syntactic, since in other places Quine himself uses semantical notions in formulating the criterion, e. g. when he says that to be assumed as an entity is to be reckoned as the value of a variable. Indeed; but every syntactically bound variable need not have any value. A quantifier binding a variable syntactically can be interpreted substitutionally as well as objectually. Variables bound by

\[ Azzouni \text{ has an argument of his own for taking the criterion to be syntactic. He argues that Quine's criterion is supposed to apply to discourses regardless of whether we take them as true, and to the extent that interpretation or translation of a discourse requires agreement, a syntactic criterion suits the demand perfectly. However, I do not see any reasons to think that interpretation would require agreement to any significant extent. The view that it would require it seems to derive from a semantic holism based on verificationism such as I have already argued (along with such philosophers as Fodor and LePore) to be absurd. Therefore the criterion can be semantic and yet apply to discourses we take to be false. } \]
a substitutional quantifier need not have any values; as Quine himself has argued (in [Qui69, page 106]), we can well quantify substitutionally over parentheses, but parentheses obviously need not have any semantical values. Substitutionally quantified variables must have substitution instances, but this is a very different thing from semantical values even in the case of variables which have both. For instance, when we write that some man is mortal, \( (\exists x) (\text{Man}(x) \land \text{Mortal}(x)) \), then Napoleon is a value of the quantified variable, but the name "Napoleon" or some other symbol for Napoleon is its substitution instance. These are very different things; Napoleon is a man but the name "Napoleon" or other symbol for Napoleon is not a man.

As e. g. Ruth Barcan Marcus has argued in [Mar72] and Dale Gottlieb in [Got76], we are ontologically committed to what we quantify over only if the quantification we use is objectual and not substitutional. Marcus and Gottlieb hope to use this insight to escape apparent ontological commitments by arguing that the quantification used in natural languages or sciences could be interpreted as substitutional; however, it is separate question whether this is possible.

Marcus’s and Gottlieb’s suggested addition to Quine’s criterion does not really constitute a revision of Quine’s theory but rather a clarification of it; Quine himself often made clear (e. g. in his reply to Ruth Barcan Marcus in [Qui66b, pages 180-182]) that he took quantification to be objectual. He also made clear in [Qui69] that he did not think substitutional quantification could serve as an account of being. However, such followers of

\footnote{Gottlieb tries to show that quantification over numbers is substitutional. However, he admits in [Got76, page 643, fn. 10] that on his interpretation of arithmetical sentences some standardly false atomic sentences may come out true, depending upon the size of the universe. This clearly implies that he does not give the ontological commitments of arithmetical theories which are commonly held to be true, but rather replaces the standard arithmetical theories with new theories which have very different ontological commitments. We may of course ask whether the new theory is preferable to the old one, and if it is then we should follow its ontological commitments rather than those of the old theory. However, showing that the new theory is better would not be simple, but would require showing (at the very least) that it is sufficient for the needs of science. Also even if such a new theory would be sufficient for the purposes it is supposed to serve, yet we could ask whether we could not replace our existing theories of arithmetic with new ones without changing quantification to substitutional. it might be that a theory in which quantification remained objectual but the quantified sentences were different would be just as good a replacement for old theories. Gottlieb would also have to show that this was not so in order to show that exchanging objectual quantification for substitutional is of any use in the philosophy of mathematics, something he certainly does not do in the article in question.}

309
Quine’s metaontology as Peter van Inwagen have occasionally suggested (e.

g. in [vI01, pages 32-36] and in [vI04, page 124]) that the whole notion of

substitutional quantification would be meaningless or unintelligible. In this

van Inwagen differs from Quine, who held (in [Qui66b, page 180]) that inter-

preting quantifiers as substitutional is an intelligible reinterpretation. In this I

think Quine is right and van Inwagen is wrong; while substitutional quantifica-

tion and other kinds of non-objectual quantification are arguably less signifi-

cant and less fundamental than objectual quantification, they are also mean-

ingful. Furthermore, there are reasons to think that the quanti-

fication used in natural languages is at least sometimes (though certainly

not always and probably not very often) substitutional. As Saul Kripke

says in [Kri76, page 375], attacking Davidsonian opponents of substitutional

quantification, we can combine substitutional quantification with referential

(i. e. objectual) in a single system. Naturally we can also combine ob-

jectual quantification with other kinds of non-objectual quantification. I

will later show that doing so will help to address some of the oldest and

most bothersome, though perhaps not most profound ontological problems,

solving them or perhaps rather banishing them as pseudo-problems (a kind

of solution I do not think to work in the case of most ancient ontological

problems, most of which I think are quite genuine).

Though the restriction to objectual quantification was implicit from the

beginning in Quine’s theory, someone might yet worry that this restriction

---

45 Many other philosophers agree with Kripke that we can combine objectual and sub-

stitutional quantification into a single system. As we have already seen Dorothy Grover

in the course of developing a prosentential theory of truth in [Gro92, §22.3, page 252]

develops a language $M_2$ in which there are both substitutional and domain-and-values

(i. e. objectual) quantifiers. Grover uses also propositional quantifiers and thinks that

they can always be substitutional; I have argued that she is wrong in this and that this

is a stumbling block for her whole theory of truth. However, with respect to individual

quantifiers she admits both substitutional and objectual quantifiers. She says, however,

something weird about their relationship in [Gro92, §2.4, page 255]; she says that

one can be defined in terms of the other, but not conversely; but neither is

"required" in giving the truth-conditions of the other. Nor is either philo-

sophically more basic.

If one concept can be defined in terms of the other but not conversely, then surely in

such cases the concept that cannot be defined in terms of the other is conceptually more

basic, according to a very common definition of conceptual basicness. Since as she agrees

substitutional quantifiers can be defined in terms of objectual ones but not conversely, this

suffices to make objectual quantification more basic. Of course, if objectual quantifiers

were required in giving the truth-conditions of substitutional quantifiers (as many other

philosophers have argued) this would make them more basic in a stronger sense; however,

this is not required to make them more basic in a weak sense.

---

310
makes Quine’s criterion of ontological commitment circular. Can objectual quantification be distinguished from other kinds of quantification without appealing to ontological commitment and saying that it is the kind of quantification which involves ontological commitment? If it could not then the criterion would certainly be circular; we would define ontological commitment with the aid of objectual quantification and objectual quantification with the aid of ontological commitment. However, I will argue that there is no real circularity. Objectual quantification can be distinguished from substitutional by the fact that objectually quantified statements and substitutionally quantified statements have different truth-conditions. Objectually quantified statements have the truth-conditions that quantified statements are given in Tarski’s theory of truth, and these do not involve any appeal to the notion of ontological commitment. The objectuality of existential quantifiers can be easily defined for standard extensional languages.

I will in what follows be using $x, y, z, \ldots, x_1, x_2, \ldots, y_1, y_2, \ldots$ as variables in the language of the first level (the object language) and $\alpha, \beta, \gamma, \ldots, \alpha_1, \alpha_2, \ldots, \beta_1, \beta_2, \ldots$ as metavariables, i.e. variables of the language of the second level (the metalanguage), except in some special cases. Such special cases will be variables of the second level language ranging over possible worlds ($w, v, u, \ldots, w_1, w_2, \ldots, v_1, v_2, \ldots$) and functions - such as assignments - ($f, g, h, \ldots, f_1, f_2, \ldots, g_1, g_2, \ldots$) and relations ($R, S, \ldots, R_1, R_2, \ldots$) and models $\mathcal{M}, \mathcal{N}, \ldots, \mathcal{M}_1, \mathcal{M}_2, \ldots$)\footnote{Actually the kind of quantification employed here is absolutely general unrestricted quantification, so the “individual” variables are interpreted as variables ranging over everything, not just over some special ontological category of individuals, and thus they can also have as values possible worlds or functions; however, specific variables will be used when I am quantifying restrictedly over world or functions.}

**Metadefinition 1** An (unrestricted) existential quantifier ($\exists_1 x$) is objectual if and only if for any formula $\phi$ and any assignment $g$, ($\exists_1 x)\phi$ is true with respect to $g$ if and only if there is such an $\alpha$ that $\phi$ is true with respect to $g(\alpha/x)$, the assignment that is otherwise the same as $g$ except that $g(\alpha/x)(x) = \alpha$.

Since universal quantifiers can be defined with the aid of existential ones, this definition naturally also determines when a universal quantifier is objectual; a universal quantifier is objectual iff the existential quantifier with whose aid it is defined in the standard way is objectual.
The purpose of this definition may become clearer if we define exactly a quantifier which is not objectual. The only non-objectual quantifier that is generally recognized is the substitutional quantifier. In order to define substitutional quantifiers we must, as Kripke has shown, select a class of expressions in the object language called the substitution class; in order to avoid circularity this class must not itself have as a member any expression containing an occurrence of the substitutional quantifier to be defined. While it is possible to take even brackets as members of the substitution class, this would be useless, so I will suppose that all the members of the substitution class are well-formed terms, simple or complex. An existential quantifier \((\exists_2 x)\) is substitutional and not objectual iff for any formula \(\phi\) and any assignment \(g\), \((\exists_2 x)\phi\) is true with respect to \(g\) iff for some term \(a\), \(\phi(a/x)\), the formula derived by replacing every free occurrence of \(x\) in \(\phi\) with the term \(a\) belonging to a substitution class \(C\), is true with respect to the very same assignment \(g\). It must be noted that as can be seen from these definitions, there can be quantifiers which are not objectual but are not substitutional either.\(^{47}\) \(^{48}\)

\(^{47}\)In fact the whole relativisation to assignments or sequences is unnecessary if the language in which we are working has only substitutional quantifiers, and therefore Tarski’s most original idea, satisfaction, also turns out to be wholly useless. Therefore if substitutional quantifiers alone would be satisfactory for logic and general semantics, then all of Tarski’s elaborate manoeuvres in his theory of truth would be unnecessary, and we would end up with a theory of truth very different from the Tarskian one. Here we see one reason, why the viability of the whole Quine-Church theory of ontological commitment is dependent on the viability of Tarski’s theory of truth in a very radical way.

\(^{48}\)Someone may yet see a problem here. Tarski’s truth-conditions only work if the quantification in the metatheory is objectual, so someone might argue that this already leads to circularity. Of course, in order to make sure that the quantification in the metatheory is objectual we can demand in the metatheory of this metatheory that its sentences have the truth-conditions given in Tarski’s theory of truth. However, we can reasonably doubt that by proceeding like this we would be led to a vicious infinite regress, so this is certainly not an adequate reply to the accusation of circularity.

However, if this objection were correct, then Tarski’s whole theory of truth would involve an infinite regress and hence be circular. Obviously the same point holds for other logical constants; for example, Tarski’s truth-conditions for disjunction only work if the disjunction in the metalanguage is ordinary classical disjunction. It is usually supposed by most philosophers that Tarski’s whole theory is not viciously circular.

In fact we may in principle escape the apparent circularity in the definition of ontological commitment by a slightly complicated procedure of elimination, though this procedure is of course of no direct practical use. Let us suppose we have - as Tarski’s truth-theory supposes (though there are of course problems about what the supposition really amounts to) - infinite hierarchies of interpreted languages and theories such that the truth-conditions of the sentences belonging to any language in a hierarchy are always given correctly by a metatheory formulated in a metalanguage immediately above it in the hierarchy. Now if all the quantifiers in some language in some such hierarchy do not have Tarskian truth-
However, this raises the question of what is common to all non-objectual quantifiers and more generally what is common to all objectual and substitutional quantifiers - why are they all quantifiers? Syntactically, on the level of formation rules, quantifiers are operators which take a variable and a formula and form another formula, but surely this is not enough for an operator to be a quantifier. One might refer to inference rules common to both kinds of quantifiers, such as introduction and elimination rules. However, since such inference rules cannot distinguish one quantifier from another, as they cannot distinguish objectual quantifiers from substitutional ones, it is doubtful if they could distinguish quantifiers from other expressions. Indeed I will argue that there are quantifiers that do not satisfy the usual introduction and elimination rules but only variants of them. Surely the notion of a quantifier can be defined semantically as well as syntactically.

I suggest that the basic idea is that a non-objectual quantifier is an operator that does implicitly quantify over some entities, but not those conditions but some more complex ones according to a correct metatheory formulated in the language immediately above it, such as substitutional ones, then none of the quantifiers belonging to a language below this language in the hierarchy are objectual either (though they might in some cases be equivalent to objectual ones). If some of the quantifiers in a language have Tarskian truth-conditions and others do not, then we can consider a sublanguage got by eliminating all sentences containing the quantifiers whose truth-conditions are Tarskian. In this way we can eliminate non-objectual quantifiers and the remaining quantifiers are objectual and carry ontological commitment.

Of course, this solution presupposes that we have available a notion of correct interpretation, and deeper problems of circularity will inevitably arise with respect to this. Basically, the answer for such doubts seems to be that a Tarskian semantical theory is not trying to define concepts like objectual quantification or classical disjunction (or, if Field is right, the reference of primitive expressions) in such a way that someone who had no previous acquaintance with them could come to understand them. It does clarify these concepts in the passing, but this is just a byproduct (though a very important one for ontology, since it involves clarifying the very notion of being which is the central notion of ontology). Such a theory must suppose that we already in some obscure way understand these concepts, and so have what a hermeneutic theory would call preunderstanding. It is surely intuitively very plausible that everyone who can use any natural language knows, even if obscurely, what it is to quantify over such entities as trees or rocks and not only what it is to quantify over linguistic expressions. What a Tarskian semantical theory is primarily trying to do is to use this preunderstanding to define the semantic values of complex expressions - concepts like truth, ontological commitment and the satisfaction and reference of complex expressions. Our understanding of the semantic values of expressions belonging to the highest level language of all those we use must then always remain at the level of a mere preunderstanding and cannot be adequately clarified. Of course this question needs further discussion, since some philosophers such as radical constructivists and formalists will not accept the supposition that we understand even obscurely he meaning of such expressions as the classical disjunction or objectual quantification, but the place of such discussion is not in this dissertation, since such discussion will best take place in works concerned with the more fundamental questions of Tarski’s theory of semantics.
which it seems to quantify over on the basis of its syntactic form. This can be understood so that a non-objectual quantifier alters systematically (but only temporarily) the interpretation of the non-logical constants and syntactical operations that occur in its scope. Most especially, predication will mean different things in the scope of such non-objectual quantifiers than elsewhere, because of the intimate connection between existence or quantification and predication that I already argued for against Lambertus Marie de Rijk.

This makes non-objectual quantifiers intensional operators; just as Frege thought that expressions within the scope of an intensional operator denoted their usual sense instead of their usual denotatum, so expressions within the scope of a non-intensional operator denote something other than their usual denotatum. However, this that they denote can be something other than their sense, as e. g. in the scope of substitutional quantifiers expressions denote themselves; because of this non-objectual quantifiers can be hyper-intensional operators which create a hyperintensional context (or scope). If quantification into intensional contexts is to be possible, it is natural to think that there must also be quantifiers that are also intensional operators, i. e. non-objectual quantifiers.

**Metadefinition 2.** An operator \((\exists_i x)\) is an (unrestricted) non-objectual existential quantifier if and only if there is such an interpretation \(J\) that for all formulas \(\phi\) and all assignments \(g\), \((\exists_i x)\phi\) is true with respect to \(g\) if and only if there is such an entity \(\alpha\) that \(\phi\) is true with respect to the interpretation \(J\) and the assignment \(g(\alpha/x)\) and for some formula \(\phi\) \(\phi\) is true with respect to \(J\) and the assignment \(g(\alpha/x)\) but is not true with respect to \(g(\alpha/x)\) and every atomic formula \(p\) is true with respect to \(J\) and the assignment \(g(\alpha/x)\) iff it is true with respect to \(I\) and the assignment \(g(\alpha/x)\).

Substitutional quantification is clearly a special case of this definition; in that case singular constants within the scope of an existential quantifier

---

49 This could also be understood in model-theoretic terms so that expressions in the scope of a non-objectual quantifier are interpreted with respect to a model even if other expressions are not, or if the other expressions are interpreted with respect to a model, then expressions in the scope of a non-objectual quantifier are interpreted with respect to a different model. However, if we want to define an unrestricted non-objectual quantifier, then we cannot use the standard concept of a model with a domain, but must instead relativise the meaning of formulas within the scope of the non-objectual quantifier solely to interpretations; restrictions on the range of quantification can be expressed as restrictions on the extensions of predicates.
denote themselves instead of the entities they usually denote (i.e. they have what scholastics called material supposition) while a predicate term denotes such a set of sequences of terms that a sequence of terms belongs to it if the sequence of their usual denotata belongs to the usual denotatum of the predicate. The novel interpretation function $J$ interprets expressions so that if $I$ is the ordinary interpretation (which is part of the object language since the object language is here supposed to be a totally interpreted language) then if $a$ is a singular term, then $J(a) = a$, and for all $n$-place predicate constants $R$ and all entities $\alpha_1, \ldots, \alpha_n \in J(R)$ iff $\alpha_1, \ldots, \alpha_n$ are expressions of the object language and $\langle I(\alpha_1), \ldots, I(\alpha_n) \rangle \in I(R)$. I will later show that other important types of arguably non-objectual quantification such as Meinongian quantification and perspectival quantification can also be viewed as special cases of this definition, which will serve to give some inductive support for the adequacy of the definition.

We can now define what is common to all quantifiers; an operator $(\exists_i x)$ is an existential quantifier if and only if there is such an interpretation $J$ that for all formulas $\phi$ and all assignments $g$, $(\exists_i x)\phi$ is true with respect to $g$ if and only if there is such an entity $\alpha$ belonging to the domain of the model that $\phi$ is true with respect to interpretation $J$ and the assignment $g(\alpha/x)$.

This criterion of objectuality can be generalized to more complex languages than the languages studied by Tarski himself. In intensional languages such as are generally studied in modern logic, the truth of sentences is relative to some indices, such as possible worlds, times, contexts etc. An existential quantifier $(\exists_i x)$ in an intensional language is objectual if and only if for any formula $\phi$ and any indices $i_1, \ldots, i_n$ and assignment $g$, $(\exists_i x)\phi$ is true with respect to indices $i_1, \ldots, i_n$ and assignment $g$ if and only if there is such an $\alpha$ that $\phi$ is true with respect to indices $i_1, \ldots, i_n$ and $g(\alpha/x)$, the assignment that is otherwise the same as $g$ except that $g(x) = \alpha$.\footnote{Other definitions generalise similarly. An existential quantifier $(\exists_i x)$ in an intensional language is substitutional if and only if for any formula $\phi$ and any indices $i_1, \ldots, i_n$ and assignment $g$, $(\exists_i x)\phi$ is true with respect to indices $i_1, \ldots, i_n$ and assignment $g$ if and only if for some term $a$, $\phi(a/x)$, the formula derived by replacing every free occurrence of $x$ in $\phi$ with the term $a$, is true with respect to indices $i_1, \ldots, i_n$ and $g$. An operator $(\exists_i x)$ is an (unrestricted) non-objectual existential quantifier if and only if there is such an interpretation $J$ that for all formulas $\phi$ and any indices $i_1, \ldots, i_n$ and all assignments $g$, $(\exists_i x)\phi$ is true with respect to $g$ if and only if there is such an entity $\alpha$ that $\phi$ is true with respect to the interpretation $J$ and any indices $i_1, \ldots, i_n$ and the assignment $g(\alpha/x)$ and for some formula $\phi$ $\phi$ is true with respect to $J$ and any indices $i_1, \ldots, i_n$ and the assignment $g(\alpha/x)$ but is not true.}
The criterion of objectuality can also be generalized to restricted quantification, though this is less important from the ontological point of view, since only unrestricted objectual quantifiers carry explicit ontological commitments (in the absolute sense; as we will see later, a quantifier restricted to the domain of a model carries ontological commitments relative to that model). However, since sentences with restricted existential quantifiers obviously imply sentences with unrestricted existential quantifiers (since if there is such an entity \( \alpha \) belonging to a set \( D \) that \( \phi \), then there is such an entity \( \alpha \) that \( \phi \), and if there is an entity \( \alpha \) satisfying the formula \( \psi \) such that \( \phi \), there is such an entity \( \alpha \) that \( \phi \)), sentences with restricted objectual quantifiers carry very obvious implicit ontological commitments, so the definition of objectuality also matters very much for ontology in the case of restricted quantifiers. A restricted objectual quantifier does not range over all entities but only over the entities contained in some set \( D \) or the sets satisfying some formula \( \phi \). This set may depend on indices - e.g. in Kripke’s version of possible world semantics the domain of quantification varies from world to world - and it may even depend on assignments if we are to represent endophoric context-dependence as in such systems as dynamic predicate logic - so I will use the symbol \( D(i_1, \ldots, i_n, g) \) for the set of entities existing with respect to the indices \( i_1, \ldots, i_n \) and the assignment \( g \). A restricted existential quantifier \( (\exists x) \) is objectual if and only if for any formula \( \phi \) and any indices \( i_1, \ldots, i_n \) and assignment \( g \) there is such a set \( D(i_1, \ldots, i_n, g) \) that \( (\exists x)\phi \) is true with respect to these indices \( i_1, \ldots, i_n \) and the assignment \( g \) if and only if there is such a \( \alpha \) in the set \( D(i_1, \ldots, i_n, g) \) that \( \phi \) is true with respect to indices \( i_1, \ldots, i_n \) and \( g(\alpha/x) \).

It is also easy to extend the definition to languages with several kinds of variables such as type-theoretic languages or many-sorted languages; in these cases the values of restricted quantifiers ranging over variables of type \( a \) (\( x_a, y_a, \ldots \)) must belong to a set \( D_a \) (which may depend on indices and assignments so we will speak of \( D_a(i_1, \ldots, i_n, g) \)). A restricted and typed existential quantifier \( (\exists x_a) \) is objectual if and only if for any formula \( \phi \) and any indices \( i_1, \ldots, i_n \) and assignment \( g \) there is such a set \( D_a(i_1, \ldots, i_n, g) \) of type \( a \) that \( (\exists x_a)\phi \) is true with respect to these indices \( i_1, \ldots, i_n \) and assignment \( g \) if and only if there is such a \( \alpha \) in the set \( D_a(i_1, \ldots, i_n, g) \) that \( \phi \) is true with respect to indices \( i_1, \ldots, i_n \) and \( g(\alpha/x_a) \), the assignment that is otherwise the same as \( g \) except that \( g(x_a) = \alpha \).
5.9.1 An Argument for Free Predicate Logic

One reason why the distinction between objectual and non-objectual quantification is important is that neglect of this distinction plays a role in the famous debate between Quine and Carnap with respect to metaontology. This debate is generally held to have been won by Quine. However, many philosophers, such as Huw Price (in articles such as [Pri09] which I have already referred to), have questioned whether Quine really had valid arguments against Carnap’s view and even how far Quine’s and Carnap’s views were in conflict. I will admit that these philosophers are in part right, as the particular arguments used by Quine against Carnap e. g. in [Qui51] are not always impressive, and there are similarities between Carnap’s and Quine’s positions. However, I will argue that there are many unclarities and fallacies in Carnap’s arguments, and that Quine’s theory provides us with material for locating them, even if Quine himself did not wholly succeed in doing so. Furthermore I will argue that so far as Quine’s view was similar to Carnap’s it was itself incorrect.

Even great philosophers have made the error of supposing that substitutional quantification would according to Quine involve ontological commitment. I think that Rudolf Carnap’s objections to Quine’s criterion of ontological commitment are partly based on this misunderstanding, a misunderstanding comprehensible in view of Quine’s own lack of clarity.

Carnap asked in [Car50] how the acceptance of new kinds of entities is represented in the language. Carnap remarked explicitly that Quine was the first to recognise the importance of the introduction of variables as indicating the acceptance of entities. Carnap thought, however, that such an introduction of new ways of speaking does not imply any assertion of reality. However, the reason for this appears to be that Carnap thought of quantification as substitutional.

Carnap said, for example, that the statement ”There are numbers.” says no more than that the new system of entities is not empty and that this is immediately seen from the rule which states that words like ”five” are substitutable for the new variables. Clearly only substitutionally quantified statements are such that their truth would follow immediately from a rule of substitution. The truth of objectually quantified sentences cannot be established this easily! The existence of numbers did not follow just from a rule of substitution in the logistic systems of Frege or Russell and Whitehead nor
does it follow from a mere rule of substitution in foundations of mathematics based on ZF, but requires several controversial axioms, such as the Axiom of Extensionality, Axiom Schema of Separation and the Power Set Axiom. Quine would have agreed with Carnap that such substitutionally quantified statements do not imply any assertion of reality; however, he would have said (or at least should have said) that such statements do not amount to the acceptance of new types of entities either.

What this suggests is that the logic of objectual quantification is strictly speaking\(^{51}\) not standard predicate logic but free predicate logic (or at least closer to free predicate logic than to standard predicate logic)\(^{52}\).

In a way Quine himself subscribed to free logic and in a way he did not. Karel Lambert says in [Lam08, page 162] that a free logician unlike Russell believes proper names and definite descriptions to be singular terms, but unlike Frege does not take all of them to denote something. Quine notoriously tended to follow Russell’s view, as he did not take such expressions to be primitive singular terms, but analysed them with the aid of quantifiers like Russell. As Lambert says in [Lam08, page 155], some of the logical systems propounded by Quine did not contain constant singular terms at all, so free logic is an expansion of such systems. However, the theorems containing singular constant symbols as defined expressions at which Quine arrives are the same as those taken as axioms in free logic, more specifically negative free logic (though not universally free logic). Surely this similarity between Quine’s logics and free logic is at least from Quine’s perspective more important than the dissimilarity; what symbols are taken as primitives is for him a rather conventional matter, but what sentences are taken as true is not. Carnap’s logic, however, was closer to standard predicate logic than to free logic. I will argue that this is the most important reason for the different attitudes of Quine and Carnap to ontology.

Since the logic of objectual quantification is not standard predicate logic but closer to free predicate logic, the ordinary rule of existential quantification, \(\phi \rightarrow (\exists x)\phi(x/a)\) (where \(x\) does not occur free in \(\phi\)), must be replaced

\(^{51}\)If as I have argued in Section 3.1.4 of this work logical conventionalism is untenable, then the question of whether standard or free logic is correct is not a conventional matter.

\(^{52}\)See [Nol11] for a brief introduction to free logic. In my view the logic of both restricted and unrestricted objectual quantification is free logic. The semantics given in introductions to free logic like [Nol11], however, only apply to (semantically) restricted free logic, since they presuppose that the quantifier must range over some specified domain.
with the following:

**Axiom 4** \( \phi \land (\exists x)(x = a) \rightarrow (\exists x)\phi(x/a) \) (where \( x \) does not occur free in \( \phi \)).

There are three main kinds of free logic, negative, positive and neuter free logic. I cannot discuss here in detail which of them is correct, though the choice between them is important for the theory of ontological commitment. According to negative free logic every atomic sentence implies existential generalizations, and therefore negative free logic implies that every atomic sentence carries ontological commitments (i.e. negative free logic contains such axioms as \( R(t_1, \ldots, t_n) \rightarrow (\exists x_i)(x_i = t_i), \) where \( 1 \leq i \leq n \)). This is probably closest to Quine’s own views.

There are some other axioms common to all forms of free logic:

**Axiom 5** \( \phi \rightarrow (\forall x)\phi \).

**Axiom 6** \( (\forall x)(\phi \rightarrow \psi) \rightarrow ((\forall x)\phi \rightarrow (\forall x)\psi) \).

**Axiom 7** \( (\forall x)\phi \), where \( \phi(t/x) \) is an axiom.

Carnap’s article appears to contain a further misinterpretation of Quine’s theory (which however is excused by unclarity on the part of Quine). The mere introduction of new variables, or any new expressions as well-formed expressions does not yet involve a person in ontological commitment in Quine’s view, as Carnap assumes, only the assertion of a sentence in which such variables occur bound by an existential quantifier does. The mere fact that one uses a language in which a sentence of the form \( (\exists i)x_i \) \( M \) is well-formed does not involve one in any ontological commitments to such entities that \( M \) even if the quantifier \( (\exists i)x_i \) is objectual. Only the assertion of the

---

53 We must in any case have some rules which allow us to infer existentially quantified sentences from atomic sentences, if we are to be justified in asserting existentially quantified sentences at all. One option is that these are formal rules applicable to every atomic sentence; this option leads to negative free logic. In this case such predicates as “fictional” or “imagined”, which are syntactically atomic in natural languages, must be treated as complex predicates in predicate logical regimentations of natural languages. Another option is that such rules are material meaning postulates applying only to atomic sentences whose predicates are some specific predicates, often called existence-involving predicates. Quine, of course, would not have liked this option because he was suspicious of analyticity and hence of meaning postulates in general. I will, however, leave here open the difficult question which of these options is better.
sentence \((\exists x)M\) (where \((\exists x)\) is objectual) as part of one’s theory - whether as an axiom of a theorem, if the theory is axiomatized - involves one in ontological commitments to such entities that \(M\). Carnap then commits the same mistake here as we have seen Gustav Bergmann did.

5.9.2 Carnap’s Relativistic Distinction between Internal and External Questions

The error pointed above undermines also Carnap’s famous arguments for a distinction between internal and external questions. This is an important point, since this distinction is one of the most important means used to attack metaphysics today, so I must dwell on it for a bit.

Carnap defined these two kinds of questions with the aid of the concept of a linguistic framework. Carnap’s definitions of these concepts are obscure in many ways, which is unusual for a philosopher like Carnap who is usually very clear - perhaps his hatred for metaphysics clouded his reasoning. This obscurity has infected innumerable, perhaps even most, metaontological debates of current philosophy with obscurity. Carnap called a system of new ways of speaking, subject to rules a linguistic framework. Carnap divided frameworks into logical and factual ones.

The first and basic obscurity in Carnap’s definition of the notion of a framework and of the two kinds of questions is that it is not at all clear what kind of rules Carnap meant. There are different kinds of rules that might be relevant here; the rules could be understood as syntactic, semantic or epistemic rules, and if they were syntactic rules they could be formation rules or inference rules. A framework could contain only rules of one of these kinds or could contain rules of several of these kinds; Carnap does not say explicitly which, but his examples suggest that rules belonging to every one of these kinds must be contained in at least some frameworks and that the different kinds of rules cannot in most cases be separated. Many theories Carnap had earlier held lead one to confuse these two kinds of rules - e. g. verificationism leads one to confuse semantic and epistemic rules - so it is likely that Carnap did confuse these different kinds of rules. However, I

\[\text{Nor77}\]

\[\text{Carnap did use the word "new", but it is hard to see what he could mean by it here, since he gave as an example of a framework the world of things, and he himself noted that we have all accepted the thing language early in our lives as a matter of course. Therefore neither the thing language or the world of things are new in any obvious sense.}\]
will argue that we must separate these different kinds of rules and that all of them cannot be viewed as constitutive of the meaning of the expressions that are used according to them, so that people capable of using expressions meaningfully can yet rationally question and debate the validity of some of these rules.

Carnap mentioned rules of confirmation and disconfirmation in connection with the world of things (which he considered a factual framework). He did not explicitly say that these rules were part of the framework of things, but if he did not think they were it is hard to see why he would have mentioned them. However, if verificationism is false, as I have already argued at length its history shows it to be, then surely we can speak about material things - make meaningful statements concerning them - without having rules of confirmation or disconfirmation, so long as we have semantic rules for expressions purportedly referring to material things. Rules of confirmation are only needed if we want to have knowledge or justified beliefs regarding material things. A person might be capable of speaking meaningfully about some entities without having justified beliefs concerning them; the later is far more difficult than the former!

Carnap also mentioned in connection with the system of natural numbers (a logical framework) customary deductive rules. The only concrete example of a rule he gives are rules of substitution, which could be understood as either formation rules or deductive rules. However, surely formation rules and deductive rules alone without any kind of interpretation do not suffice to speak about anything, even about numbers. If these are the only kind of rules Carnap considered, then he had just reverted to the position of [Car37], whose falsity he had already recognized in [Car48]. In order to speak about any entities, we need also semantic rules for our expressions, such as ostensive definitions, but not rules of confirmation.

Since Carnap's theory is a relativistic theory, it will be helpful to consider it in the light of relativism more generally. Panu Raatikainen has argued in [Raa01] quite plausibly that Putnam's mereological argument for relativism was based on a confusion between languages and theories. Other philosophers have noticed a similar confusion in relativistic theories more generally; e. g. Edward Pols does so in [Pol92, page 5]56. I will argue that

---

56Pols sees such a confusion in too many theories, claiming that the majority of English-speaking academic philosophers are guilty of such a confusion. I do not see where he has
the arguments of at least Huw Price and probably Carnap himself are based on a similar confusion (though Quine may not have always escaped such a confusion either; Chomsky accuses him of such a confusion in [Cho69, pages 53-54]). In this case this may be a confusion that is at the root of relativism more generally.

Huw Price defends Carnap’s distinction between internal and external questions in [Pri67]. He distinguishes six ingredients of Carnap’s position. The third of them is that frameworks are ontologically committing. Price indeed goes further and says that the very use of concepts is ontologically committing. Price thinks that Quine would have accepted this ingredient. This is indeed plausibly an ingredient of Carnap’s position (though Carnap did not like the word ”ontological” due to his anti-metaphysical prejudice). However, many commentators think that a linguistic framework for Carnap was scarcely anything more than a language (as e. g. Susan Haack suggests in [Haa76, page 458] and as Matti Eklund says in [Ekl09]), and Carnap does e. g. speak of the thing language and the framework of things as if they were the same or at least inseparable. However, according to at least most57 of Quine’s formulations it is theories, not languages, that carried ontological commitments. Therefore if this common interpretation is right Price is wrong in saying that frameworks carry commitments and Quine should not have admitted it.

Price has then confused languages and theories just like Raatikainen argued Putnam did, and if Price’s interpretation of Carnap is so far correct, Carnap also fell earlier prey to the same confusion. Even if it is not clear whether Quine himself would have admitted it, yet if we follow Church’s theory of ontological commitment, it is clear that frameworks need not be ontologically committing. As I argue that Church’s theory of ontological commitment is better than Quine’s original, this would suffice to refute Price’s defence of ontological relativism. At least those frameworks whose deductive rules are those of free predicate logic do not carry any ontological commitments. I think that this is the crucial false assumption whose grounds for so sweeping an accusation and would not want to join in it.

57Quine’s formulations in [Qui66a] may to some extent excuse Price’s misunderstanding. In this article Quine does remark in [Qui66a, page 65] that to say that to say that there is such an entity as roundness is to say that from a context ‘... roundness...’ we may infer ‘(∃x)(...x...).’ However, Quine’s formulations of his criterion in [Qui53d] imply that we say that there is such an entity as roundness only by actually saying ‘(∃x)(...x...).’
falsity completely undermines Carnap’s and Price’s arguments against metaphysics.

Raatikainen says in [Raa01, footnote 22], however, that he is not assuming free logic, and argues that Putnam’s argument can be shown invalid even in standard logic. It is not clear whether he would have any objections to the use of free logic as such. It might well be that Putnam’s argument can be resisted even within standard logic, but Carnap’s and Price’s arguments could not be.

Later metaontologists have modified Carnap’s definitions in various ways, usually resulting in still greater obscurity. The phrases ”linguistic framework” and ”conceptual scheme” are often used as synonyms. I have argued that Carnap did not make it clear whether his frameworks are languages or theories, though they are more likely to have been languages. While many later relativists have carried on the obscurity, many of them have often understood a framework or scheme explicitly as a complete theory. Quine himself used the phrase ”conceptual scheme” to stand for theories, and indeed pretty big theories (which may explain why he did not see that Carnap’s frameworks were languages, and so did not carry ontological commitments). This seems to be a rather different conception, though this is not sure because Carnap does not make clear just what kinds of rules a framework may contain. After all, inference rules can generally be transformed to axioms in notational reformulations of an axiomatic theory, so a system of inference rules could be taken to correspond to a weak theory, the purely logical theory derived from those axioms by the inference rules corresponding to them. However, it seem to me a plausible interpretation of Carnap’s view that a framework in Carnap’s sense need not contain any more statements than these logical truths (which correspond to rules of deduction) and possibly other analytic truths (since Carnap did not at this stage of his development distinguish clearly between logical truths and other analytical truths, though he was to do so soon after in [Car52]).

However, a framework or scheme in the broader sense of a theory can contain all kinds of non-logical and even synthetic statements. Therefore if this interpretation of Carnap is correct, the two concepts differ profoundly. Speaking of theories as frameworks makes conceptual relativism into the obviously absurd claim - though a claim we have seen was already implicit in Neurath and Hempel - that we cannot speak about anything except relations.
between theories, that we cannot speak about anything non-linguistic at all.

Carnap called questions of the existence of new entities within the framework internal questions. However, Carnap gave two seemingly quite different characterizations for external questions; he said initially that internal questions are questions concerning the existence or reality of the system of entities as a whole, but later he said that an internal question is a question posed prior to the acceptance of a new framework. I will argue that a question concerning the existence or reality of the system of entities as a whole can be a question within a framework and indeed given many ways of interpreting what a framework is has to be such. Carnap claimed later in the same article that a question concerning the existence of a system of entities as a whole such as "Are there numbers?" can be interpreted as either an internal or external question, while his initial characterization implied that it must be interpreted as an external question.

Carnap argued that such an assertion as "There are numbers." is trivial and hence not metaphysical if taken as an internal question while if it is taken as an external question it is not a theoretical question at all but a practical question. One way of countering this argument would be to reject the whole distinction between theoretical and practical questions, as the pragmatists would have done. Even Quine seems to have hinted at such a rejection at the end of [Qui51]. However, I will argue that we can reject Carnap's distinction between internal and external questions even while accepting the distinction between theoretical and practical questions.

I will show that this argument only seems plausible because Carnap runs
together different notions of a framework, which contain different kinds of rules or different rules. The basic fallacy in Carnap's argumentation results from the fact that his notion of a framework is equivocal because as we have shown he does not make it clear what kind of rules are contained within it and what rules of each kind have to be contained in it. Therefore depending on how the notion of a framework is understood, it may be either not the case that we need accept a framework in order to talk about new entities or it may not be the case that the acceptance of a framework would lead by itself to the acceptance of any new system of entities and therefore would make questions about the existence of such entities trivial.

If it is supposed to be essential to a framework that it contains the deductive rules of standard predicate logic, then we do not need a framework in order to speak about any entities. Indeed, we do not need any deductive (or epistemic) rules at all to speak about any entities in the sense of making (even if only irrationally) such statements that terms occurring in them denote such entities if they exist. However, if talking is understood to entail arguing, we may need them to justify the claims we make about them, since we need to argue in order to justify claims, even if wholly hypothetical ones, and argumentation requires deductive rules. However, even in order to argue hypothetically concerning new entities, we do not need a framework with the rules of standard predicate logic, for we can use the deductive rules of free predicate logic instead. Therefore in order to have arguments relating to the new entities it suffices to use a framework in a sense in which a framework can include only the deductive rules of free predicate logic, and no existential claim at all taken as a claim within such a framework is trivial.

Let us examine the two options Carnap supposes the metaphysician to have in turn.

Carnap supposes that a metaphysician cannot take metaphysical questions to be trivial and hence not internal. However, there is a problem here even independent of the problems with Carnap's notion of a framework. Carnap seems to be just have accepted without question Kant's view that metaphysical truths must be synthetic, which we have shown to be based on a misunderstanding of traditional metaphysics. However, I have already argued that a mildly deflationary conception of metaphysics is possible which takes at least many of the traditional metaphysical questions to be analytical ones and hence relatively trivial (though this would not make them any
more trivial than mathematical questions on the traditional view of analytic philosophy) and that such a conception has many important historical precedents.

However, perhaps more importantly, Carnap’s argument does not work even against an inflationary conception of metaphysics, and this is because of the problems with Carnap’s notion of a framework. Traditional metaphysical questions can be understood as internal questions without being taken to be trivial relative to a framework, if the kind of framework used is weak enough, e. g. if it contains only semantical rules or besides them only the deductive rules of free predicate logic.

It is possible that all the new predicate terms have the empty set as their extension and if the logic of the language to which the new expressions are added is free predicate logic it is possible that all of the new terms are non-denoting. Indeed, if also the logic of the metatheory is free it is even possible that predicates not only have an empty set as their extension but have no extension at all. The new rules themselves may lead us to suppose that none of the new entities exist, producing only such statements as \( \neg(\exists x)P(x) \), \( \neg(\exists x)Q(x) \), \( \neg(\exists x)(x = a) \), \( \neg(\exists x)(x = b) \) . . . Even a (formulation of a) complete theory answering every question which can be asked within such a framework might in the case of some frameworks contain only such negations of existential generalizations and nothing else. Also even if a framework has to contain confirmation and disconfirmation rules, such rules can lead a person to disconfirm every statement which claims that an entity belonging to the system of entities in question exists. Therefore none of the internal questions of existence need be trivial as Carnap claims.

So an extreme idealist who thinks that material things do not exist (but allows the statement that they do to be meaningful) could surely, contrary to what Carnap presupposes, accept the thing language - so long as he does not use standard predicate logic but instead free predicate logic - and use it without accepting the world of things. He could use it just to say of all possible physical things that they do not exist. Indeed, such an idealist has to use the thing language if he is to deny the existence of physical things. Such an immaterialist could accept that whatever analytic truths are held to hold of physical objects, e. g. ”Physical objects have a location.”, ”Physical objects are extended.”, etc. are vacuously true though (or rather since) no physical objects exist.
We have then seen that if the notion of framework is weak enough, metaphysical questions can be taken as internal questions even without being taken to be completely trivial. However, if the notion of framework used is strong enough (e.g., if it contains the deductive rules of standard predicate logic), then traditional metaphysical questions can be understood as external questions while still being taken to be theoretical questions.

It is not transparent what Carnap means or could mean by a question prior to the acceptance of a framework, which is one of his characterizations for an external question. If it is supposed to be a question posed without the acceptance of any framework, and a framework were taken to consist only of formation rules and semantic rules used in order to form sentences, then external questions are of course impossible. It is obviously impossible to pose any question without using any expressions. It is equally impossible to pose a question regarding some purported entities without using expressions purportedly denoting those entities, by either using uninterpreted expressions or expressions denoting only other purported entities. However, there is no reason to think that any metaphysicians would have attempted to perform such an impossible feat.

It is indeed possible to pose questions without having any axiomatized theory and hence without deductive rules, so it may be we do not need the exact kind of framework Carnap conceived of. However, in order to try to answer such a question we do need to use (though not mention) some deductive rules, but they can be those of free predicate logic.

However, an external question could also be understood in other ways. It could be a question asked in a metalanguage (using a different framework than the framework about which we ask) about whether the expressions contained in a framework designate anything (as already seen by Haack in [Haa76, page 460]). Since the aim of Carnap’s article is to defend Tarskian semantics, it follows from the very aim of the article that Carnap’s frameworks are indeed very similar to Tarski’s interpreted languages, if the rules Carnap speaks about can be understood as semantic rules, which provide interpretation for the expressions in a framework; at least many commentators such as Haack in [Haa76, page 458] presume that a framework is an interpreted language. If this is correct we could speak about object frameworks and metaframeworks just as we speak of object languages and metalanguages, and say that we ask questions about the object framework within a metaframework. We can also compare different object frameworks within a metaframework, and in many cases say that one framework is better than another in the sense that it has fewer non-denoting terms and its terms denote more entities.
nap has to accept that such a question is meaningful and theoretical. Such questions are naturally relevant for metaphysics; however, they are rather metametaphysical than metaphysical questions.

I have argued that the most important difference between Carnap and Quine relevant to ontology was their different theories of quantification. However, it is often assumed instead that the most important difference was elsewhere, most often that it was in their theories of analyticity. Carnap says that such a claim as that numbers exist (and presumably any question within what he would call a logical framework) is analytic. Quine himself saw in [Qui51, page 71] the basic point of contention between Carnap and himself to lie in Carnap presupposing a distinction between analytic and synthetic statements. This may indeed be a good reason for those who reject the distinction between analytic and synthetic statements to also reject the distinction between internal and external questions in the case of logical frameworks or even Carnap’s whole distinction between factual and logical frameworks. However, I do not think Quine yet located the crucial error in Carnap’s argumentation. The rejection of a distinction between analytic and synthetic would not do anything to undermine Carnap’s distinction of internal and external questions with respect to factual frameworks, while the rejection of standard predicate logic does. Furthermore, even those who accept a distinction between analytic and synthetic statements do not have to accept the distinction between internal and external questions. They should be dubious about using the notion of analyticity in the way Carnap uses it here.

One reason for this can be seen by reminding ourselves of the fact that many analytical philosophers and early logical positivists generally held that analytic statements could not be existential. E. g. Ayer says in [Aye36a, page 128] that to assert that an object exists is always to assert a synthetic proposition. Because of this they could not have accepted the way Carnap uses the notion of analyticity here though they would have seen nothing wrong in the notion of analyticity itself (though curiously Ayer later did accept the distinction between internal and external questions in [Aye67, pages 49,71], perhaps not seeing its contradiction with his view of analyticity or changing his view of analyticity).

Analytical philosophers who are not positivists do sometimes think that existential claims could be analytic - a logicist who does not agree with
Russell’s view that logical truths are synthetic has to make this claim, as does the theory of such modern logicians as Cocchiarella or Hale and Wright - but they would not have think that such claims would have followed just from a rule of substitution, but required instead less trivial axioms like the Axiom of Comprehension or the Axiom Schema of Separation. However, even those who accept that the claim that abstract entities such as numbers - or the propositional functions or classes to which they are reduced - exist is analytic would hardly hold the same of the claim that concrete entities, such as physical objects (or things as Carnap calls them) or minds or even sense-data, exist. It is just absurd to say that "Physical objects exist." is an analytic sentence. It is still more absurd to say that it would be a conventionally true sentence in any sense but that in which all sentences are conventionally true because of the conventionality of linguistic symbolism (which Carnap’s position also implies according to some interpreters), and therefore such a sentence cannot be trivial at all, no matter how interpreted. If standard predicate logic implies that such a claim could be analytic, this is a reason not to use standard logic in serious philosophical discussions of semantics, and instead replace it with free predicate logic.

Since Quine was not a type-theorist, he of course did not think that the introduction of new kinds of entities would require the use of new kinds of variables, only of new predicates and sentences asserting them to be non-empty; however, while this difference is more pertinent to their different views relating to metaphysics, I do not think this either really captures the crucial difference between Quine and Carnap in this question.

Neither Carnap nor Quine explicitly made a distinction between standard and free predicate logic or between objectual and substitutional quantification in the course of their debate, and I think this is a reason why their debate never reached any conclusion and why neither of them is completely correct. Carnap is partly right in claiming that a nominalist could accept a framework which contains predicate variables. A nominalist need not think that the statement that universals exist is meaningless - while such a claim is indeed nominalist it is only one variety of nominalism. A nominalist can

\[\text{Quine apparently sees nothing odd in the claim that such a sentence as "There are physical objects." would be analytic; apparently since he has convinced himself that the concept is bogus, the analyticity of any sentence is just as implausible to him as that of any other. However, Norton sees in [Nor77, page 80] that such a claim is very odd and problematic for Carnap himself.}\]
instead think that the realist’s view is just false, whether analytically false or even synthetically false. Such a nominalist can accept the use of a framework whose formation rules allow the use of predicate variables bound by objectual quantifiers if the deductive rules of that framework are those of free higher-order predicate logic instead of standard higher-order predicate logic (or if the framework has no deductive rules at all). A nominalist can accept such a framework since he can use it to assert a theory with such theorems as \( \neg(\exists P)(\exists Q)(P = Q) \), where \( P \) and \( Q \) are predicate variables, as such a theorem claims that there are no universals. However, contrary to what Carnap implicitly claims and in accordance with Quine’s view, a nominalist cannot accept the use of a framework whose formation rules allow predicate variables as well-formed expressions, whose quantifiers are objectual and whose deductive rules are those of standard higher-order predicate logic, since such framework would force him to admit the truth of the sentence \( (\exists P)(\exists Q)(P = Q) \), where \( (\exists P) \) is an objectual quantifier. Of course, a nominalist could accept standard predicate logic if the quantifiers occurring in its theorems were all interpreted substitutionally.

Ironically Quine’s own later thesis of ontological relativity in [Qui68] is rather similar to Carnap’s distinction between internal and external questions. This raises serious doubts about the consistency of Quine’s thesis of ontological relativity with his criterion of ontological commitment - doubts which I unfortunately do not have room to pursue here\(^{61}\).

### 5.9.3 Ontological Commitment and the Metaphysics of Modality and the Semantics of Propositional Attitudes

If we admit the legitimacy of intensional metatheoretical notions such as possible worlds, contrary to Quine’s own views, we must notice that there may also be other varieties of quantification besides objectual and substitutional quantification. This is important for it is one of the factors which will enable us to apply our theory to controversial questions in the metaphysics and metametaphysics of modality.

The definition of objectuality in possible world semantics is just a special

\(^{61}\)Quine’s ontological relativism is closely connected with his structuralism (as presented e. g. in [Qui08a], as his arguments for both make heavy use of proxy functions. However, because of this it seems to me that they are both vulnerable to Newman’s old argument against structuralism in [New28]. Quine’s argument for ontological relativism is also made suspect by the use of behaviourism in it, as I have already argued.
In intensional logic it is usual to interpret quantifiers by quantifying in
the metalanguage over individual concepts or world-lines, functions (or par-
tial functions) from possible worlds to individuals (or more generally, since
such quantification can also bind predicate variables, to entities). In this ap-
proach assignments associate individual variables with individual concepts
or world-lines, never with individuals, but the predicates are taken to typ-
ically have the values of the world-lines in the worlds, not the world-lines
themselves, in their extension. It is not at first sight obvious if this kind
of quantification is objectual; however, I will argue that it is not, or at least
it is not objectual quantification over individuals (if individuals are taken
as entities which are not functions or sets or properties, but are rather Ure-
lements in the set-theoretical sense). At least such quantification is not
unrestricted quantification, and this already bars it from carrying explicit
objectual commitments. However, it is not even restricted quantification re-
stricted to individuals, but at most quantification restricted to world-lines.

More exactly, in the simplest versions of this kind of theory, \((\exists x)\phi\) is true with respect
to a world \(w\) and assignment \(g\) if and only if there is such a world-line \(f\) (some theories add
that \(f\) must belong to some set \(F\)) that \(\phi\) is true with respect to \(w\) and \(g(f/x)\). However,
if \(\phi\) is e. g. a monadic atomic formula \(P(x)\), then \(P(x)\) is true with respect to \(w\) and \(g\)
iff \(g(x)(w) \in \|P\|_{w,g}\), not iff \(g(x) \in \|P\|_{g,w}\), as it would be if we were straightforwardly
quantifying over world-lines.
However, it is not presented as second-order quantification over world-lines, but as first-order quantification over individuals. From this we can argue further that we are not quantifying straightforwardly over world-lines any more than over individuals, and hence are not quantifying objectually over anything at all. This may not seem sufficiently justified yet, so I will add further evidence later, relying on the commonly accepted connection between quantification and identity. First, however, I must give examples of such non-standard interpretations of quantification.

For instance, Anil Gupta has (following and simplifying the ideas of Aldo Bressan [Bre72]) used such quantification in his theory of common nouns and natural kinds [Gup80]63. Independently of Gupta and Bressan, Jaakko Hintikka and logicians influenced by him (e. g. Ilkka Niiniluoto in [Nii82]) have developed a logic of perception in which a peculiar variety of this kind of quantification (a variety where world-lines are relativised to subjects of propositional attitudes such as perception) was called perceptual quantification. This concept was later generalised to the notion of perspectival quantification so that it would apply to the logic of other propositional attitudes besides perception, for instance to the logic of belief or the logic of imagination. I will speak of perspectival quantification, since the differences between perception and other attitudes are not relevant to the questions I ask in this work, which do not relate to the logic of propositional attitudes or to phenomenology (to which that logic might be applied) but to general ontology. Niiniluoto defines the general concept of world-lines as follows in [Nii82, page 119]: If $W$ is the clas of possible worlds, and if $Dm(w)$ is the class of individuals which exist in a world $w \in W$, then, for $\emptyset \neq V \subseteq W$, a function $f : V \rightarrow \bigcup_{w \in W} Dm(w)$ such that $f(w) \in Dm(w)$ for each $w \in V$ is a world-line on $V$.

I want to argue that this kind of quantification need not carry any explicit ontological commitment any more than substitutional quantification. This is no doubt a highly controversial statement. Philosophers who make use of this kind of quantification have expressed different view on this question.

---

63When we combine the theory of Gupta with my theory, it suggests that nearly all of the quantification employed in natural languages is not objectual, and hence that all of the ontological commitments in natural language sentences are implicit, which may seem highly surprising. However, Quine himself warned against naive attempts to apply his criterion directly to natural languages, so this is not a sufficient reason to reject either theory, but just serves as an additional reason for precaution in using the criterion.
Hintikka himself thought that this was just ordinary objectual quantification. He says in [Hin75, page 32]:

Thus no unusual sense of quantification is being assumed here, certainly nothing remotely like that bete noire of Quine’s, substitutional quantification. Rather, what we face here are the consequences of an objectual interpretation of quantification in a situation where one is concerned with possible worlds.

However, other philosophers who have developed Hintikka’s logic of propositional attitudes further or just studied it have held different opinions. Føllesdal in [Føl69, page 178] argued that, given Quine’s arguments against quantified modal logic, Hintikka’s semantics only makes sense if the values of the variables are taken to be expressions. Following him Tuomo Aho held (see [Aho94, pages 231-234]) that all quantifiers should be interpreted as substitutional in attitude contexts. Aho, unlike Føllesdal, even indicates that this is in accordance with Hintikka’s intentions, saying:

Already Hintikka’s original model sets were defined in such a way that the occurrence of $(Ex)\phi$ in a model set also guaranteed that the set contained $\phi(t)$ for some singular term $t$ . . .

However, both Føllesdal’s and Aho’s arguments have serious problems. I think that Quine’s arguments have been refuted many times and in many different ways (as they have several apparent weaknesses), even by Hintikka himself in later articles, though nothing is ever definitely settled in philosophy, so of course the debate is still going on. Føllesdal himself admits in his article that there are several ways to avoid Quine’s nihilistic conclusions, most prominently [Føl69, page 183] eliminating definite descriptions contextually (or just distinguishing between definite descriptions and logically proper names - however you do it (even Kripke’s way of taking all syntactically proper names to be logically proper will get you far) - and allowing substitution only of logically proper names in attitudinal contexts)

Of course this still leaves some similar problems concerning substitution in propositional attitude contexts, which have been discovered afterwards, such as Kripke’s puzzles about belief. However, these problems were not proposed by Quine and it can be doubted how deeply they are analogous to Quine’s problems. Anyway, I think that they are sufficiently dissimilar that I will not have to deal with these tangled messes of muddled debate here when asking the only tangentially related question which kinds of quantification in propositional attitude contexts are objectual.

64Of course this still leaves some similar problems concerning substitution in propositional attitude contexts, which have been discovered afterwards, such as Kripke’s puzzles about belief. However, these problems were not proposed by Quine and it can be doubted how deeply they are analogous to Quine’s problems. Anyway, I think that they are sufficiently dissimilar that I will not have to deal with these tangled messes of muddled debate here when asking the only tangentially related question which kinds of quantification in propositional attitude contexts are objectual.
Actually the solution Hintikka offers to the problems is similar to the one Church offered which Føllesdal examines along with Hintikka’s. As he says in [Føl69, page 177], in Church’s system variables take intensions as values. If Føllesdal’s interpretation of Church’s theory were correct, then the quantification in Church’s theory would be just a plain variety of restricted objectual quantification. Hintikka’s system is here in many ways a more sophisticated development of Church’s theory, since as we already hinted world-lines are really intensions or individual concepts, only relativized to subjects of propositional attitudes. According to Føllesdal Church’s system is not really non-extensional, and because of this it avoids Quine’s problems. For the sake of consistency he should have said the same about Hintikka’s system; if Church’s system is not non-extensional - note that I am not agreeing that Føllesdal would be right - then Hintikka’s is not either and of course also avoids Quine’s problems just like Church’s does. Then the quantification in Hintikka’s system would be just restricted objectual quantification as well as in Church’s system. However, I will argue that in fact neither Church’s nor Hintikka’s quantification is objectual in attitudinal and other non-extensional contexts.

The problem with Aho’s argument is, of course, that Hintikka’s later definitions of model sets (or model structures to use more modern terminology) were not any more defined in this way so that the occurrence of $(Ex)\phi$ in a model set also guaranteed that the set contained $\phi(t)$ for some singular term $t$, and neither are the model structures Aho himself uses in the same work where he says this. The values of variables can hardly be both expressions and world-lines; even if models were constituted out of linguistic expressions, and a world-line were just a function to linguistic terms, yet a function whose value was a linguistic term would be different from a linguistic term. However, since those early days both Hintikka and others have used a more general kind of models, and surely in most models world-lines are not expressions or functions to expressions. As Kripke remarks, substitutional quantification is unproblematic in modal contexts; if

---

65In other ways Hintikka’s system is a simplified version of Church’s theory, since where Church’s theory is a higher-order type-theoretical theory, most formulations of Hintikka’s logic of propositional attitudes are first-order. However, Hintikka has also more complex formulations where he uses independence-friendly quantifiers which make his system located between first-order and higher-order systems, so it is very hard to accurately characterize the relation between Hintikka’s and Quine’s systems in such a case.
quantification were substitutional, then there would be no need for the very elaborate apparatus of world-lines.

Therefore Aho’s view does not seem to be defensible. However, Hintikka’s own view can also be doubted. While Hintikka is clearly right in claiming that his quantification is not substitutional, he may not be right in his further claim that it is just objectual quantification either. It must be admitted, however, that that question of what quantification is objectual is in part stipulative. The distinction between objectual and substitutional quantification was originally drawn in extensional languages, and there need be no unique way to extend it to stronger languages, such as intensional and especially modal languages. Nevertheless, I will argue that contrary to what Hintikka suggests, the introduction of world-lines is not the necessary consequence of an objectual interpretation of quantification in a situation where one is concerned with possible worlds. The way I have defined objectual quantification seems to me to be simpler. If objectuality is defined as I have defined it, then the kinds of quantification introduced by Hintikka and his followers are neither substitutional nor objectual.

They satisfy my definition of non-objectual quantification, Metadefinition 2. If the ordinary interpretation is $I$, then the new interpretation introduced by the perspectival quantifier is such a function $J$ that individual constants are interpreted as their ordinary intensions, i.e. if $a$ is an individual constant, then $J(a, w) = \{\langle w, I(a, w) \rangle : w \in W\}$ and a predicate is interpreted so that it takes as arguments the values of its arguments in the world i.e. for n-place predicates $R$ and all world-lines $\alpha_1, \ldots, \alpha_n$ it holds that $\langle \alpha_1, \ldots, \alpha_n \rangle \in J(R, w)$ iff $\langle \alpha(w), \ldots, \alpha_n(w) \rangle \in I(R, w)$. This interpretation leads to the same truth-values in all purely extensional contexts, but yields different truth-values for sentences containing intensional operators like propositional attitude ascriptions, as it allows quantification to be over world-lines and not ordinary individuals, so that the same variable can pick out different individuals from different worlds.

In any case, the vital question is not how to use the word “objectual”. The vital question is what kinds of quantification carry ontological commitment, and we have reasons to doubt whether the kinds of quantification introduced by Hintikka need carry ontological commitment. The answer to this question depends on our definition of ontological commitment. Of course, this may also contain a stipulative element, though any reasonable
answer must also agree with the long tradition of ontological inquiry.

If we are right that perspectival quantification is not objectual, perspectivally individuated objects do not really exist and there is no need for a theory of categories to supply them with a category of their own. This is of course no defect in the theory of perspectival quantification, but rather a merit of it. This is not a novel conclusion either; some of Hintikka’s followers have already come to it. For example, Tuomo Aho writes in [Aho94, page 264]:

it would be possible to call these perspectival objects even intentional objects. But, and this is a great merit of the world-line method, the intentional objects clearly do not form a a new category of existing individuals in the world.

However, Aho does not explicitly distinguish objectual and non-objectual quantification so I do not think that he yet had sufficient reasons for this conclusion; Aho actually claims that all quantification in attitudinal contexts is substitutional, and hence implicitly non-objectual, so I do not know what criterion he could offer for deciding which objects form a category of existing individuals and which do not. Surely the fact that perspectival objects exist in other possible worlds is not enough to bar all of them from the category of existing individuals, since it is clear that many perspectival world-lines often pick out individuals from the actual world and so many perspectival objects also exist in the actual world. The theory of this dissertation allows us to given an exact reason for Aho’s intuitively correct conclusion.

However, it can be argued that all statements and theories which make use of such quantification carry implicit ontological commitments, since they imply statements expressible in terms of objectual quantifiers; how-

\[\text{Actually, in Hintikka’s original logic every world-line picks out something from the actual world (or the world counted as actual), namely that object which is the cause of the perception. This is not so in Aho’s development of the logic, and I think that this constitutes a fruitful generalisation over Hintikka’s logic; however, it does not suffice to give Aho the materials to construct a criterion for which objects belong to the category of existing individuals.}\]

\[\text{Many perspectivally quantified sentences carry implicit commitments at least in the sense I have already defined in this section, material implicit ontological commitments. They may also carry formal implicit ontological commitments, though this depends on the controversial question whether logics of propositional attitudes in which perspectival quantifiers are used count as genuine formal logics, i. e. whether the relation of implication investigated in them is genuine formal implication.}\]

\[\text{E. g. to give a trivial example, if } (\exists x) \text{ is the perspectival and } (Ex) \text{ the objec-}\]
ever, such implicit commitments require extensive semantical analysis to ferret out. Even if some weak statements or theories employing perspectival quantification would not carry even implicit ontological commitment according to some metatheories, yet when a person uses perspectival quantification which is identified as such, i. e. a person uses quantified statements to which someone correctly gives a semantics making use of individual concepts and possibilia, then the one who gives such semantics to it would in any case be committed to such individual concepts and possibilia, since the metatheory he uses carries such commitments. Therefore the use of perspectival quantification ultimately involves (at least in some way) even greater ontological commitments than ordinary objectual quantification.

There is a simple test which in many cases suffices to show whether quantification in an intensional language is objectual or not; a quantifier \( \forall_i x \) is objectual only if it satisfies the formula \( \Box_1 \forall_i x \forall_i y (x = y \rightarrow \Box_1 (x = y)) \), where \( \Box_1 \) is the operator of logical necessity. Any quantifier whose truth-conditions are straightforward extensions of Tarski’s into an intensional language must satisfy this formula. Neither substitutional quantification nor perspectival quantification satisfy this condition. Substitutional quantification does not pass this test in general because substitutional quantifiers can have as substitution instances terms which are not rigid designators, but denote different entities in different worlds. A variety of substitutional quantifiers, \((\exists x)(P(x) \& P_a(Q(x)))\) implies the sentence \((Ex)(P(x))\), since if the perspectival world-line picks out something which satisfies the formula \(P(x)\) from the actual world, then there must exist something in the actual world which satisfies this formula, and since it exists in the actual world, it must exist absolutely. Because of this, the sentence \((\exists x)(P(x) \& P_a(Q(x)))\) carries implicit ontological commitment to entities which satisfy the formula \(P(x)\). To give a less trivial example, we must make heavy use of the possibly controversial device of extending the object language to find out the ontological commitments expressible in the metalanguage, as introduced in Metadefinition 8, and introduce such doubtlessly controversial machinery as higher-order quantifiers or explicit set-theoretical expressions to the enriched object language. If we use higher-order quantifiers, then the sentence \((\exists x)(\forall(P(x)))\) implies the sentence \((EQ)(\Box((Ax)(Ay)(Q(x) \& Q(y) \rightarrow x = y)) \& \Box((Ex)(Q(x) \& P(x))))\) in any language which contains besides the perspectival individual existential quantifier \((\exists x)\) the objectual individual existential quantifier \((Ex)\) and the objectual predicate quantifier \((EQ)\) and the objectual individual universal quantifier \((Ax)\). The implied sentence quantifies explicitly over world-lines with the aid of the predicate quantifier \((EQ)\), if we interpret predicate terms as functions from worlds to sets of individuals, as is usually done; the condition \(\Box((Ax)(Ay)(Q(x) \& Q(y) \rightarrow x = y))\) implies that the property designated by the predicate variable \(Q\) relative to some assignment is only instantiated by one individual in every world, and therefore it is a function and hence is or at least corresponds to a world-line in the broadest sense, though of course the sentence does not say anything about what kind of world-line it is or corresponds to.
quantifiers whose instances were exclusively rigidly designating terms would pass the test, and such quantifiers would indeed not differ significantly from objectual quantifiers restricted to a set composed of the entities those rigid terms designated. Even such substitutitional quantifiers would indeed not be objectual, as they would not have any explicit ontological commitments, but they would be rather obviously equivalent to objectual quantifiers, since statements in which they occur would have the same implicit ontological commitments as statements employing objectual quantifiers. However, in order to be able to know or say that a substitutitional quantifier is such that all of its instances are rigidly designating we must already be capable of using objectual quantifiers, so objectual quantifiers cannot be reduced to substitutional ones by just using a language all of whose terms are rigidly designating.

That perspectival quantifiers were not objectual was not surprising. More surprisingly, however, it seems that the physical quantifiers, which Hintikka and his followers contrast with perceptual or perspectival ones, are not objectual either. We might indeed argue that physical world-lines are rigid and hence physical quantification is equivalent to objectual; however, doubts must be raised about the rigidity of physical quantification. According to Hintikka and many of his followers, world-lines can merge and split; this means that perceptually or physically individuated entities which are different with respect to one world can be the same with respect to another world. Formally, if \( f \) and \( g \) are world-lines and \( w \) and \( v \) are worlds, it is possible that \( f(w) = g(v) \) but \( f(v) \neq g(v) \). It follows from this that a quantifier \( (\forall_i x) \) defined with the aid of such world-lines cannot fulfill the formula \( \Box_1(\forall_i x)(\forall_i y)(x = y \rightarrow \Box_1(x = y)) \) and therefore it is not objectual. This means that we cannot speak about the absolute identity of physically individuated objects, but only of their identity with respect to worlds. However, since existence and identity go together, as Quine saw, this implies that where there is no absolute identity there cannot be absolute existence either.

\[69\] According to Aho [Aho94, page 264]

The difference of physical and perspectival world-lines is not the same as that between things in themselves and phenomena.

This is indeed correct, but not because perspectively identified objects would not be phenomenal in the widest sense - they are - but only because physically identified objects are also phenomenal, which is quite in accordance with the semantic Kantianism Hintikka has professed.
and therefore physically or perspectivally individuated objects cannot have any place in a rational ontology.

It must be remembered that this condition is only necessary, not sufficient for objectual quantification; quantification over rigid world-lines (world-lines which pick out the same individual from all possible worlds) also satisfies this test, so it is not always easy to find out if we are dealing with objectual quantification or not. However, while such quantification over rigid world-lines does not carry any explicit ontological commitments, it yet comes so close to doing so that the difference does not matter for most purposes, since such quantification is obviously equivalent to objectual quantification in the sense that it can be very easily translated into objectual quantification and has the same implicit ontological commitments as objectual quantification.

Hintikka has actually doubted whether the notion of transworld identity makes sense independent of a conceptual system, so he would also doubt the intelligibility of the whole notion of objectual quantification. Other philosophers carrying on his work have denied the intelligibility of absolute transworld identity quite explicitly, e. g. Tero Tulenheimo in [Tul09]. This would naturally also raise doubts about the intelligibility of any kind of absolute notion of ontological commitment. Such doubts would of course (despite the notorious antagonism between Quine’s and Hintikka’s philosophies) fit well together with Quine’s own later theory of ontological relativity, which might be held to support such doubts; however, such doubts threaten to trivialise the whole theory of ontological commitment. However, the use of the technique of perspectival world-lines does not imply this kind of relativism, and others who use it are free of such relativism. We may even argue that since the notions of absolute identity and objectual quantification in any case usually are used and arguably must (if we are to avoid infinite regress) be used in the metalanguage, it is just self-defeating and incoherent to deny their meaningfulness. Let us just look at Niiniluoto’s definition of world-lines quoted above; the domains of worlds, $Dm(w)$, for all possible worlds $w \in W$, are supposed to exist before the definition of world-lines; the definition of world-lines presupposes their existence. However, since the metalanguage is naturally equipped with quantifiers and the identity predicate, this implies that we can formulate such statements in the metalanguage as $(\exists \alpha)(\exists \beta)(\exists w)(\exists v)(\alpha \in Dm(w) \land \beta \in Dm(v) \land w \in W \land v \in W \land w \neq v \land \alpha =$

\[339\]
(1) and intelligibly ask whether they are true. The metatheory Niiniluoto uses does not determine whether such statements are true or not, but it does determine that they are meaningful and determinately either true or false. Since we can speak about such identities or their lack in the metalanguage, we can also expand the object language so that we can speak of them, i.e., so that besides physical and perspectival quantification it also includes genuinely objectual quantification over individuals. Such statements, however, identify entities in different worlds prior to and independently of the world-lines, contrary to what Hintikka and most of his followers claim. If all such statements are false, we have an ontology of world-bound individuals, like those of Leibniz and Lewis. Actually, however, many of Hintikka’s statements seem to imply that at least some of such statements must be true.

5.10 The Problem of Non-existence; an answer to Meinongians

The possibility of employing non-objectual quantification alongside objectual is vital for solving the ancient problem of non-existence famously raised by Parmenides, namely the problem how (or even whether) we can speak truly about that which is not. It helps us to see that we can think and speak about what is not (for example about fictional objects) contrary to Parmenides, but that we do not yet have to admit non-existent and even impossible objects into our ontology as Meinong or at least some Meinongians thought to our ontologies contrary to Meinong.

---

70. This holds at least if we assume that the logic of the metatheory is a classical logic, i.e., two-valued and without truth-value gaps, so that it validates the law of excluded middle. Neither Hintikka nor any of his followers has ever stated that it would not be, but Hintikka has famously appealed to a logic without excluded middle in other contexts, e.g., in his theory of truth. Using a non-classical metatheory might help to avoid this consequence; however, so far as I know no one has explicitly made this kind of suggestion. The motivation for using non-classical logic in order to avoid paradoxes is entirely different from the motivation of using it to deny the intelligibility of absolute existence.

71. I will be using the word "object" for anything that can be quantified over whether objectually or non-objectually, while I will be using the words "entity" and "being" for anything that can be quantified over objectually. Therefore every entity and every being is an object but every object need not be an entity or being. Both words, "object" and "entity" are then dummy sortals (like the Greek word "αὐτό" and the Latin words "ens" or "res" that "being" and "entity" usually translate, or German equivalents such as "seiendes" or
The distinction between objectual and non-objectual quantification will also help us to counter the most important counter-arguments that have been directed at Quine’s theory of ontological commitment, for example in [Hod72]. Michael P. Hodges points out that there are cases in natural languages where quantifiers do not function in the way that Quine’s criterion requires, such as the following:

**Example 4** *There is something he is looking for, namely, the fountain of youth.*

This sentence does not seem to carry ontological commitment to a fountain of youth. One of them most influential recent critics of Quine’s theory of ontological commitment, Jody Azzouni uses as a central example in his argument against Quine’s view of ontological commitment in [Azz04, page 62] an example sentence which is concerned with quantification over fictional objects.

**Example 5** *There are fictional mice that talk.*

Hodges says correctly that it is not difficult to find contexts in which such sentences as Example 4 have a perfectly acceptable use. The same holds for Azzouni’s example.

However, this does not prove much against Quine’s theory. It is enough for a Quinean that quantifiers have some widespread contexts or readings

---

72While Azzouni rejects Quine’s view wholly, his argument is to a large part aimed specially at the claim that Quine’s criterion of ontological commitment is trivial. I do not wholly agree with Quine that the criterion would be as wholly trivial as he thought, so Azzouni’s examples may prove this much; however, I will argue that they do not show that Quine’s view, with the qualifications and clarifications I have added to it, is correct.
or uses where they function in the way his criterion requires, if he can
distinguish such readings or uses from those where they do not in some
systematic way. It would also be desirable and bolster the Quinean’s case
if he could tell (at least roughly) how quantifiers function in those other
contexts or what those other readings and uses of them are, but this is not
necessary to save the theory from counterexamples such as those of Hodges
and Azzouni.

It is suspicious that most important counter-examples concern very spe-
cific contexts, namely contexts concerned with propositional or objectual
attitudes or fictional objects, contexts that many semanticists would label
intensional. If a person looks for a fountain of youth he must believe or at
least suspect that there is a fountain of youth, so this example is concerned
with a propositional attitude, though this may not be obvious at first sight.
We can then suggest as a first approximation that quantifiers semantically
express ontological commitments (at least with respect to one reading - and
plausibly in most of them the most natural reading - of them) in all ex-
tensional contexts (though even in such contexts they may have pragmatic
uses where they pragmatically implicate something other than ontological
commitment).

The intensional contexts mentioned are of course the contexts in which
according to Frege’s famous theory of sense and reference (which Church,
whose theory of ontological commitment I am trying to develop further here,
also developed in [Chu51]) linguistic expressions have different references
and senses than usually, referring to or denoting what is is usually their
sense. In this case of course if Frege’s theory is at least close to truth,
one would expect that sentences would also in such contexts carry different
ontological commitments than usually, carrying ontological commitment to
the senses of the expressions that occur in them to whose referents they
would usually carry ontological commitments. Therefore Example 5 would
carry ontological commitment to the sense of the phrase "mice that talk”
(whatever kind of entity that is - it could be taken to be a function from
possible worlds to sets of possible talking mice, as the structured meaning
of the phrase "mice that talk", or as a sui generis entity) instead of to its
referent, i. e. it would not carry ontological commitment to talking mice.73

73 This might count as an instance of the fifth of the six possible ontological interpre-
tations of his example Azzouni lists, the cancellation option. Azzouni objects in [Azz04,
Frege’s theory remains, of course, controversial. However, all systematic semantic theories recognize that contexts of this kind are specially problematic, so using them as counterexamples to a theory of how any type of expression such as quantifiers generally function does not carry conviction. Also the most prominent alternative to Frege’s theory as a general semantic theory trying to account not only for extensional but also for intensional contexts (such as modal contexts) has been Kripke’s theory of direct reference as presented in [Kri72]. However, it seems to me that at least one of the most prominent current opponents of Quine’s theory of ontological commitment, Jody Azzouni, cannot be happy with Kripke’s theory, since Kripke treats fictional entities as existing abstract entities (as clearly stated in [Kri11, page 63] and elaborated in [Kri13]), a claim that Azzouni finds in [Azz04, page 72] to be repugnant to common sense. However, if Azzouni’s interpretation of his example sentence does not fit in with either of the two dominant systematic semantic theories (and he does not give any systematic semantic theory of his own either or refer to a third theory) this is surely a reason to suspect that his interpretation of his example is not likely to be correct. Of course, this does not refute it conclusively, but then Azzouni admits that his arguments do not establish his interpretation conclusively either, so I think that this suffices to turn the burden of proof away from Quine’s theory.

Quine’s criterion concerns primarily the quantifiers of predicate logic and of artificial languages we can build from it by adding to it non-logical
constants, not the quantifiers of natural languages. Natural language quantifier phrases like "there is" do often, perhaps even usually, function as the quantifiers of predicate logic do. However, in order to fully address Hodges’s objection, an adherent of Quine’s criterion must give some account of those circumstances in which natural language quantifiers do function differently.

The distinction between objectual and non-objectual quantifiers I develop in this work is the key to one such account; quantifiers that occur in (the most natural readings of) example sentences like Example 4 or Example 5 are non-objectual ones, and Quine’s criterion only applies to objectual quantifiers. We can say that something exists really only if it is quantified over objectually in a true statement. However, it is desirable to develop this reply further; it is not fully satisfactory to dismiss such quantifiers as non-objectual, but we ought to say something what kind of non-objectual quantifiers they are (though as I said I do not think that this is absolutely necessary for a defence of Quine’s criterion).

I will quickly sketch how some of these notorious problems can be solved in my theory. I cannot go into details, and there are no doubt more difficult versions of these problems which require a more subtle approach; especially, this kind of approach may not be able to get rid of commitment to possibilia, though it will help us to get rid of the more clearly absurd commitment to Meinongian impossible objects. In order to avoid ontological commitment to possibilia we would have to get rid of objectual possibilistic quantification in every theory we use; it is not enough to have all quantifiers in the object language be actualist but employ possibilist quantification in the metatheory, as actualists commonly do when dealing with modal logic. There certainly could have existed objects for which we do not have names, so it is hopeless to try to analyse possibilist quantification as substitutional quantification. However, whether possibilist quantification can be analysed as some other kind of non-objectual (non-entitative) quantification - e. g. analysing possibilia in terms of maximal sets of compatible properties or of individual concepts - is a very difficult question which I cannot address in this dissertation (and for which I have no ready answer at all).

Hodges’s and Azzouni’s objections are really the same as a common objection directed by Meinongians against Quine, for example by Richard Routley in [Rou82], even though Hodges might have been startled to find
out that he was expressing Meinongian ideas. Routley objects to Quine that most things do not exist, and Hodges’s fountain of youth is of course an example of the kind of things that according to Meinongians do not exist.

A Meinongian thinks that we can say truly that that some objects do not exist, since it seems that such objects as the golden mountain or the square circle or the fountain of youth do not exist. Meinong himself said in [Mei04, §3, p. 9] (see [Mei60, page 83] for an English translation) that those who like paradoxical modes of expression could well say that there are objects of which it is true that there are no such objects. However, this paradoxical

---

Azzouni explicitly opposes Meinongianism. However, it is not at all clear how Azzouni’s view would differ from Meinong’s view, especially as both views are rather subtle and can be accused of being a bit obscure. Azzouni’s own characterization of Meinong’s view seems to be inaccurate, so it seems likely that Azzouni just does not recognize the similarities between his view and Meinong’s, especially as Azzouni does not include any work by Meinong in his list of references. Azzouni does refer to works by Meinongians such as Routley and Parsons and Deutsch, but it is of course not clear that these authors have interpreted Meinong correctly or that Azzouni has understood them correctly. Azzouni says in [Azz04, page 65] that the Meinongian conclusion is that some things that are don’t exist or aren’t real, and reminds philosophers in [Azz04, page 71] that fictional objects don’t exist in any sense whatsoever. However, Meinong did not think that the object of knowledge had to be in any sense at all; see e. g. [Mei04, §3, pages 7-9], translated in [Mei60, §3, pages 81-83], where Meinong stresses that an object need not exist or subsist, where existence and subsistence are the two senses of “Sein” i. e. “being” he recognized. Meinong also agreed with Azzouni when Azzouni says (appealing to common sense) in [Azz04, page 72] that fictional objects are not abstracta, since in Meinong’s view abstracta subsisted though they did not exist but fictional objects did not even subsist. Meinong may not always be fully consistent in holding that fictional objects do not exist in any sense, but he is surely at least as consistent and arguably even more consistent than Azzouni in holding this. Azzouni contradicts himself blatantly in saying in [Azz04, pages 115,116] that “exists” does very often operate as a locally contrastive device without ontological significance. If “exists” used contrastively applies to some fictional objects, then such fictional objects surely exist in some sense, nameley the contrastive sense, even if it is a sense carrying no ontological significance. It follows from the deflationary theory of truth that Azzouni himself holds that if the word “exist” is true of, i. e. applies to, some fictional object, then that fictional object exists. On the other hand, what Azzouni says when he contrasts his theory with the theory of the Meinongian Routley in [Azz04, pages 55,56] seems more to the point. Azzouni says that non-existent objects do not have properties and that there is nothing to discover about them. Here he does seem to differ from Meinongians. However, there is a problem here. Since Azzouni is a nominalist it would seem that he cannot say that even existent things have properties. Nevertheless he insists that existing things have properties even though there are no properties, and this again sound paradoxical in the same way as Meinongian views. Does it not imply that nonexistent properties have the property of being had by existent objects? Furthermore, even where Azzouni’s theory does seem to genuinely differ from Meinongian theories, it is not clear that this is wholly in his favour. There seems to be a sense in which there is something to discover about fictional objects. E.g. a person who is only vaguely familiar with Sherlock Holmes can discover more things about him by reading more of Conan Doyle’s stories.
cal statement easily appears directly contradictory if it is formalized in the wrong way. If existence is expressed by quantifiers and if both quantifiers in such a statement are taken to be the same kind of quantifier, whether objectual or not, we get the contradictory statement \((\exists x)(\neg(\exists y)(y = x)))\). This drives many Meinongians to deny that existence would be expressed by quantifiers, and instead take it as a primitive predicate. According to them, the statement should be formalized as \((\exists x)(\neg E(x))\), where \(E\) is the existence predicate. This is indeed a consistent formalization. However, the statement can be symbolized consistently even if existence is taken to be expressed by quantifiers, so long as we use together with the objectual quantifier also a substitutional quantifier (whose substitution instances include complex terms like definite descriptions besides simple constants) or some other non-objectual quantifier. We can then formulate such a paradoxical statement as the wholly consistent statement \((\exists_2 x)(\neg(\exists_1 y)(y = x)))\), where \((\exists_1 x)\) is the unrestricted objectual quantifier and \((\exists_2 x)\) is a non-objectual quantifier. This solution is of course similar to a Meinongian solution in one respect; many Meinongians would also advise that two different quantifiers are to be used. However, the difference is that this statement will according to the theory of this dissertation not carry any ontological commitments, at least not any explicit ontological commitments (or even implicit commitments to wholly non-existent objects), since the quantifier \((\exists_2 x)\) with the broadest scope in the statement is not objectual, so we are not committing ourselves to a Meinongian ontology in using such a statement. However, this may not be contradict Meinong himself, since Meinong himself stressed in [Mei04, §2, p. 5] (see [Mei60, page 79] for an English translation) that metaphysics (which according to him has to do with everything that exists and so comes close to ontology in the sense in which I employ the word in this dissertation) cannot take on the task of the theoretical consideration of the object as such, from which it follows that theory of objects is different from metaphysics and also from ontology if ontology is considered as either identical with metaphysics or as a part of metaphysics in the way Wolff defined it. My solution only contradicts the views of some followers of Meinong or Meinongians who have misinterpreted Meinong’s philosophy or developed it in an untenable direction. It might be presented as an objection to this that

---

76Actually philosophers who describe themselves as Meinongians speak of a Meinongian ontology less often than those who oppose Meinong’s ideas, but Meinongians also do
some philosophers in the phenomenological tradition, most notably Husserl, have associated the theory of objects with ontology (see e.g. the introduction to the third Logical Investigation in [Hus70, page 435]), where the concept of ontology used is the one common in phenomenological philosophy which I described earlier, according to which ontology (the science of necessary and a priori truths concerning possible entities) is not the same as metaphysics but a more fundamental discipline on which metaphysics is based. However, unlike Meinong these phenomenologists did not usually admit impossible objects into their ontology or theory of objects even though they admitted abstract objects - both tropes and universals (many of which Meinong would also have thought to be part of the subject matter of ontology) and perhaps even possible objects. Therefore they agreed that all objects in the Meinongian sense, which includes impossible objects, are not a part of the subject matter of ontology. Therefore even a purely historical study suggests that in any case, no matter in which of the historically motivated ways metaphysics and ontology are thought to be related, ontology is not a discipline concerned with Meinongian objects. This view can be further explained and defended by arguing - contrary to many Meinongians - that the objectual quantifier is not a restricted quantifier and the non-objectual quantifier is not a logically primary expression. Instead any non-objectual quantifiers can always be defined in the metalanguage by using objectual quantifiers, so that the objectual quantifiers are more basic (and ultimately there is one unrestricted objectual quantifier which is the most basic of all sometimes lapse into speaking of Meinong's theory of objects as a Meinongian ontology or even - contrary to Meinong's own express statement - as a Meinongian metaphysics. E.g. Parsons speaks in [Par74, page 567] of Meinong's theory of objects as an unusual ontology. On the other hand, Routley says in [Rou82, page 152] that the ontological problem is a problem as to what exists or has being, so the Meinongian theory of objects is not a solution (or even part of the solution) to an ontological problem according to Routley.

Matters are complicated by questions regarding Husserl's controversial concept of a noema, which is supposed to be the object of an intentional act as such; e.g. the noema of an act of perception is the perceived as such. There is much discussion about whether noemata are objects or contents of mental acts, based of course on how such qualifiers as "as such" are to be understood. If they were taken as objects then they would be similar to Meinong's objects, so phenomenology which is according to Husserl a material ontology, and includes as a part noematic phenomenology, would then include a general theory of objects, including impossible ones. However, most commentators seem to agree that noemata are not objects, at least not in any straightforward sense. In any case, the theory of objects which include impossible objects would still not be a part of the most general kind of ontology, namely formal ontology, and it would not be an independent science as Meinong claimed but just a part of phenomenology, being correlative to the theory of noeses.
and expresses the notion of being simpliciter).

In the simplest case the non-objectual quantifier in question may even be taken as substitutional; talk about the golden mountain may be considered as talk about the expression "the golden mountain". Here of course the quantifier presupposed in the definite description must of course be understood as objectual in order to avoid circularity, as Kripke has pointed out; if the sentence is symbolized as $(\exists_2 x)(\neg (\exists_1 y)(y = x))$, then the substitution instance for the variable $x$ would have to be some definite expression like $(\iota_1 z)(G(x) \wedge M(x))$ This implies that the sentence would be derived from the sentence $\neg (\exists_1 y)(y = (\iota_1 z)(G(z) \wedge M(z)))$ and the definite description which has appeared in this form should be defined contextually in the well-known way so that the whole sentence could be reduced to the form $(\neg (\exists_1 y)((\exists_1 z)(G(z) \wedge M(z) \wedge (\forall_1 u)(G(u) \wedge M(u) \rightarrow u = z)) \wedge y = z))$.

However, such an unproblematic (if sometimes laborious) solution to the problem may not always be possible, but a subtler approach may be needed. In fact many Meinongians such as Parsons, Castañeda and Zalta have already developed machinery for developing such a theory of non-existent objects, which can easily be stripped of its apparent ontological commitments. Most of these modern developments of Meinongian theories reduce Meinongian impossible objects to sets of incompatible properties or to properties which cannot be instantiated. For instance, Terence Parsons develops in [Par74] a Meinongian semantics, which is intended as a set-theoretic reconstruction of Meinong’s theory. Parsons thinks that this reconstruction will only communicate the structure of the theory of objects and help to show its consistency, but I think that it can also be taken to communicate the material content of Meinong’s theory. We might say that Parsons gave a model for Meinong’s theory that she did not think to be its intended model while I think that it can be taken to be its intended model. In Parsons’s semantics when we apparently quantify in the object language over individual objects (using individual variables) we give the truth-conditions for such sentences by quantifying over sets of properties. Parsons thinks that such sets of properties represent Meinongian objects but are not identical with them; however, she admits in [Par74, page 578] that one might insist that objects are the sets that she uses to represent them. While I do not want to say that such sets are Meinongian objects (since I do not think that there are any Meinongian objects) I nevertheless think that Meinongian objects

348
could be reduced to such objects as logical constructions out of them.

It seems to me that in this case the kinds of quantification used in the object language sentences and in their interpretations in the metalanguage diverge so widely that the quantification in the object language must be taken as non-objectual, though it is not substitutional. Such quantification will satisfy my definition of non-objectual quantification, Metadefinition 2. The new interpretation function introduced by the non-objectual quantifier is the interpretation function that associates with a name the set of properties possessed by the referent of the name; i.e. \( J(t) = \{ P : P(I(t)) \} \), if \( I \) is the ordinary interpretation and \( t \) is an individual term and \( P \) is taken as a variable in the metalanguage ranging over properties (however properties are understood in detail, whether as primitives or as sets of possible individuals or as functions from possible worlds to sets of individuals or even very inadequately as sets of actual individuals or whatever). It will interpret a predicate \( P \) that a Meinongian would say refers to a nuclear property as \( \{ \alpha : I(P) \in \alpha \} \). However, it will associate a predicate that a Meinongian would say denotes an extranuclear property with its ordinary interpretation. Parsons admits in [Par74, page 569] that she does not have a criterion that would sort antecedently identified properties into nuclear and extranuclear, so the reinterpretation of predicates may have to proceed case by case. However, predicates that a Meinongian would classify as extranuclear would probably not be taken as atomic in any perspicuous language, as they include such properties as existence and possibility, so it may be possible to interpret any predicates as designating the set of all sets of properties containing the property they would ordinarily designate. Meinong (as interpreted by Parsons) takes monadic predication as primitive and tries to reduce polyadic predication to monadic; even if this does not work for all polyadic predications, it may work for those predications dealt with in Meinong’s theory. Therefore if I have been correct in arguing that variables bound by non-objectual quantifiers do not carry ontological commitments, then Meinong’s theory as interpreted by Parsons does not carry ontological commitments to impossible entities (even though this may be contrary to Parsons’s own intention) but only implicit ontological commitments to properties and sets of properties. While these commitments are of course themselves highly controversial, they are yet less controversial than commitments to impossible objects and are commitments that are quite ac-
ceptable if realism about properties and sets is acceptable, as I argue in this dissertation.

The same approach can in my view be applied to fictional entities or (more accurately expressed) fictional objects such as Azzouni’s counterexample concerns. Weirdly enough, realism concerning fictional objects has been very popular in recent analytical philosophy; even such famous philosophers as Saul Kripke (as seen in [Kri11]) and David Kaplan have been realists about fictional entities. So has one philosopher who is one of the most famous supporters of the Quinean criterion of ontological commitment, namely Peter van Inwagen; see [vI01, page 38]. There is a debate concerning creationism; whether fictional entities are created by human beings or exist independently of them as abstract objects such as numbers are said to exist according to Platonists. However, it seems to me that if realism is implausible anywhere at all, it is surely implausible with respect to fictional entities. You can argue that fictional entities by definition are not real, do not genuinely exist. This is Azzouni’s view (in [Azz04, page 72]), and though I disagree with his theory of ontological commitment, his intuitions seem to me to be right in this respect. Therefore I think that it would be desirable for a Quinean to find a response to the counterexamples that does not imply the existence of fictional objects. If fictional entities do not exist at all in the literal sense of the word then they are not literally created (since only something which literally exists is literally created) so creationism is false but neither are fictional entities Platonistic abstract objects (as van Inwagen was forced to think, since he could not accept the idea of non-objectual quantification). While the theory of abstract entities I defend in this dissertation shares similarities with Platonism and might even not entirely inappropriately be called Platonistic, I do not want to extend a similar approach to fictional objects. However, it is obvious that we can quantify over fictional entities as in Azzouni’s example of fictional mice. Since my theory of ontological commitment is so similar to van Inwagen’s it is important that I show that it does not compel me to accept his views on fictional objects (partly because unlike van Inwagen I accept the possibility of non-objectual quantification).

Such quantification will have to be non-objectual, if we are to be anti-realists about fictional entities, as common sense says we should be. However, it is not plausible to hold that all such quantification would be straight-
forwardly substitutional. Azzouni has a quite good argument (in [Azz04, pages 72,73]) why this would not be satisfactory; many characters in fiction are nameless. For instance, it is understood that there are many more hobbits, dwarves and elves in Middle Earth than are named in Tolkien’s works. However, we can hold that such quantification is some kind of non-substitutional and non-objectual quantification. In many cases, it might be perspectival quantification, which we have already examined. However, this will probably not work adequately in the case of inconsistent fictions, which are a very difficult case. However, inconsistent fictions are probably a very common case, since authors are fallible and often make mistakes. For example, Arthur Conan Doyle apparently said both that Doctor Watson’s first name was James and that it was John. Clearly no one can have two different first Christian names according to the naming system in use in western countries so this is an inconsistency. In some cases apparent inconsistencies can be explained away by more subtle interpretation of a piece of fiction, but there are surely many cases where this will not work.

We then have a very difficult semantic problem. However, such impossible fictional objects seem similar to Meinongian impossible objects, so it seems likely that any adequate semantic theory for statements about any impossible objects might also be extended into a semantic theory for statements about impossible fictional entities. Many Meinongian ontologists, such as Parsons, have developed a theory of fictional entities with the aid of Meinongian objects; however, if I have been correct in arguing that there is no need for a Meinongian ontology in order to have a semantical theory of impossible objects, then it should also be possible to develop a theory of fictional entities without presupposing any Meinongian ontology. Therefore I think that even in such cases statements quantifying over fictional entities might be given truth-conditions by quantifying in the metalanguage over necessarily uninstantiated properties or sets of incompatible (but possible and perhaps even actual) properties of particulars which the fictional work (such as a novel or short story or comic or film etc.) or the (actual, not fictional\textsuperscript{78} narrator of the fiction is in a semantical relation to.

\textsuperscript{78}In theories of fiction a distinction is often made between the actual narrator or narrators and the fictional narrator or narrators. E. g. in the case of the Sherlock Holmes stories, Arthur Conan Doyle is the actual narrator while Doctor Watson is the fictional narrator. Of course, since the fictional narrator or narrators is or are fictional and belongs or belong to a fiction (whether to the same fiction he narrates or to another fiction, as in more complicated cases where we have stories within stories) he or they must also be
5.11 Church’s criterion of ontological commitment and truth

Alonzo Church suggested in [Chu58] the correction to Quine’s criterion that since we should allow for models where variables do not have a non-empty range it is only existential quantification, not universal, that carries ontological commitment\footnote{Church’s suggestion does carry some risk; it might be misunderstood to imply that universally quantified statements would not be ontologically relevant. Universally quantified statements, which do not carry commitment to the existence of any entities but just deny the existence of some entities, can (so long as the universal quantifier is objectual and unrestricted) be part of ontological theories as well as statements that affirm the existence of some entities. Indeed, it can even be argued that such statements are a more crucial part of ontological theories than existentially quantified ontological statements, being typically more difficult to establish and more controversial. For example, the materialist’s assertion that all entities are material, i.e. that there are no immaterial entities, is surely an ontological assertion, though it does not carry any explicit ontological commitments to the existence of any entities. We can indeed say that the materialist’s assertion implicitly carries commitments to the existence of material entities since it presupposes that they exist, but the truly controversial part of his claim is that there are no other entities; this is what separates him from dualists. Someone might well express this by saying “universally quantified statements involve ontological commitments”, and there would be nothing necessarily wrong with this. However, this is not how the concept of ontological commitment is usually employed since the work of Quine and Church, and I think that it is best to stick to the terminology that has become standard. Even if we use the phrase “ontological commitment” solely for commitments to the existence of entities, we can use some other phrase such as “ontological relevance” or “ontological implication” to refer to commitments to either the existence or the non-existence of entities. We could divide ontological implications into explicit and implicit and into different kinds of implicit ontological implications similarly to the way we in this dissertation divide the concept of ontological commitment.}. I think Church is right here. Even more importantly, Church also tries to formulate his version of Quine’s criterion more exactly. The following is the formulation in [Chu58, page 1014]:

Schema 1 The assertion of $(\exists x)M$ carries ontological commitment to entities $x$ such that $M$.

Church says that this is to be understood as a schema in which ‘$x$’ may be replaced by any variable, ‘$x$’ may be replaced by any name of the same variable, M may be replaced by any open sentence containing no other free variable than ‘$x$’ and ’$M$’ may be replaced by any name of the same propositional form. This formulation seems to me to be a lot clearer than any of

reduced away in the metalanguage if we are to succeed in developing a non-realistic theory of fictional entities. I am of course using the word “narrator” here in a wide sense, so that the director or scriptwriter (or both) of a movie also count as narrators.
Quine's own formulations, and far clearer than most of the formulations of Quine's other followers (with the possible exception of Cartwright's rather clear formulations) and quite plausible, at least if we are speaking about explicit and not about implicit commitment. Thus it is strange that it has not figured very much in recent discussion concerning ontological commitments. Chihara is one of the few who have noticed that Church's formulation is clearer than Quine's formulations; however, as I will show, Chihara still partly misinterprets Church's formulation.

Those familiar with modern formal semantics should here experience a sense of déjà vu. Church's formulation of the criterion of ontological commitment sounds a lot like Tarski's criterion of material adequacy for the concept of truth, the famous T-schema or Convention T\(^80\) (see [Tar83a] and [Tar44, page 344]). One famous (but not completely accurate) formulation Tarski gives for the T-schema is the following:

**Schema 2** \(X\) is true if, and only if, \(p\).

Tarski explains that this is to be understood as a schema in which '\(X\)' is the name of '\(p\)'. The same principle is used both in Tarski's T-schema and in Church's criterion of ontological commitment. In both cases we have a schema where both an expression and its name occur but no other (expressions expressing) material notions except (an expression expressing the) the semantical notion that is to be defined occur. It is crucial for both conven-

---

\(^{80}\)Convention T" is more often used, but I think that "T-schema" is really a better name, for "Convention T" suggests that the schema is somehow especially conventional, but it can be argued that the schema is no more conventional than any sentence or sentential schema is, and this is just because of the conventionality of all language, whether natural or artificial. If we choose to use the word "true" as it is actually used (or in a significantly similar way, even if we clarify it to make it a technical term), we cannot consistently choose not to hold the T-schema as true. Of course, if we use the word "true" in an artificial language, then how we use it is to some extent conventional. However, the T-schema is not just a stipulation for how a semanticist chooses to use the word "true", but an explication of the way it and its translations have actually been used in natural languages for a long time. Of course, explication or regimentation does always hold a conventional element; however, it can be argued that in the case of the word "true", the conventional element is actually lesser than in most cases of explication. We can argue that the T-schema could actually serve not merely as an explication but even as the foundation of an analysis in the sense of G. E. Moore of the central meaning of the word "true" (if analysis in the Moorean sense is at all possible). Of course, it is not meant as a description of the way the word true" is (pragmatically) used as the ordinary language philosophers would desire; just as in any semantical analysis, we must exclude such peripheral and possibly metaphorical uses as in expressions like "true friend", ignore conventional implicatures etc..
tion T and Church’s formulation of the criterion of ontological commitment that the same expression is both used and mentioned in expressing them. I propose that we call Church’s formulation of the criterion Convention C and the C-schema (where the letter "C" comes either from "commitment" or from "Church" in honor to Church) in analogy to Convention T and T-schema.

Tarski’s Convention T is a biconditional. Convention C as I have formulated it is not a biconditional or other kind formula derived from other formulas with the aid of propositional functions; however, it can be translated into a conditional.

**Schema 3** An assertion $S$ carries ontological commitment to entities $x$ such that $M$ if $S$ is of the form $(\exists x)F$ where $M$ is a translation of $F$ to the metalanguage.

This can even be transformed into an equivalence if we use the narrower notion of explicit ontological commitment instead of the wider notion of ontological commitment:

**Schema 4** An assertion $S$ carries explicit ontological commitment to entities $x$ such that $M$ if and only if $S$ is of the form $(\exists x)F$ where $M$ is a translation of $F$ to the metalanguage.

Famously though Tarski held that the concept of truth was not strictly speaking applicable to natural languages, he yet illuminated it (e. g. in [Tar83a, page 343]) for pedagogical purposes with the aid of such examples taken from natural languages such as English and Polish as the following Example 6:

**Example 6** The sentence "Snow is white" is true if and only if snow is white.

This Tarski’s most commonly known formulation of the T-schema is misleading in one way. It is disquotational i. e. homophonic, which implies that the metalanguage must be an extension of the object language. However,

---

81 Tarski’s own example may not in retrospect have been as fortunate as possible, since it contains the mass term "snow" and there exist especially difficult problems concerning the semantics of mass terms quite distinct from the general problems associated with truth. Later discussion has focused more on examples like the following: The sentence "John is a doctor" is true if and only if John is a doctor.
elsewhere Tarski makes clear that a more careful formulation of the scheme should also take into account cases where 'X' is the name of a sentence whose translation 'p' is; for instance, in [Tar44, page 350] he says:

The requirement in question can be somewhat modified, for it suffices to assume that the object-language can translated into the meta-language...

Since sceptical doubts can be raised concerning the general intelligibility of the notion of translation, using this notion makes Tarski’s theory more controversial. Also, there are wider and narrower notions of translation, so this opens the possibility of developing a Tarskian theory in different directions by using different notions of translation. Clearly, a notion of translation according to which translation presupposes strict synonymy between the translated expression and the translating expression is unnecessarily strong for the purposes of a theory of truth, but a too loose notion that allows any loose paraphrases to count as translations may on the other hand be too permissive. Tarski’s less than sufficiently careful formulation of the schema has lent support to deflationist interpretations of Tarski’s theory, such as the Disquotational Theory of Truth. Künne has in [KÖ3, §4.2.1page 226] called the narrower interpretation of the T-schema the Disquotation Schema. A wider formulation of the schema that makes use of the notion of translation has also less trivial instances which are not disquotational or homophonic, since in them the metalanguage is not an extension of the object language, but the object language can be a completely different language, e.g.:

Example 7 The sentence "Schnee ist weiss" is true if and only if snow is white.

Actually, the requirement can be modified further (and ought to be, since it still sounds too deflationist). Tarski’s actual theory can apply also to languages (or fragments of languages) whose sentences cannot be translated into the metalanguage, so long as these languages are semantically open sentences and are lower in the Tarskian hierarchy of semantically open languages than the metalanguage and have similar logical forms as the sentences of the metalanguage.

Even though Quine was sceptical about whether the concept of ontological commitment is applicable to natural languages - which scepticism
Church apparently did not share - Church’s formulation can like Tarski’s be illuminated, at least for purely pedagogical purposes, with the aid of examples from natural languages. When this is done the similarity becomes even more apparent. The examples naturally have a disquotational form just like Tarski’s examples of the T-schema, though like Tarski’s T-schema the C-schema also has less trivial instances which are not disquotational.

The following is an example of an instance of Convention C, the criterion of ontological commitment, completely parallel to Tarski’s own example:

**Example 8** *The sentence "There is white snow" carries ontological commitment to white snow.*

The following example may be more illuminating since it does not contain a mass term:

**Example 9** *The sentence "There are men." carries ontological commitment to men.*

We can also give examples where the metalanguage is not an expansion of the object language, e. g. the following where the object language is Finnish and the metalanguage English:

**Example 10** *The sentence "Naisia on olemassa." carries ontological commitment to women.*

In fact it seems plausible that the criterion of ontological commitment is simply an application of some more general semantical principles implying also the Convention T and similar schemata involving other semantical concepts such as denotation, satisfaction, obedience etc. For example we can give as an example of the condition of material adequacy for the definition of the concept of obedience a schema with the example "The person addressed obeys the command "Shut the window" if and only if he shuts the window". This schema would (when developed further in a precise manner) allow us to formulate a correspondence theory of obedience. Thus we could generalise truth-conditional semantics to other sentences besides statements, namely to commands. Similar correspondence conditions could be formulated for other kinds of sentences. While the notion of denotation must probably be taken as a primitive in a Tarskian theory as Field has argued, yet we can
give as a necessary though not sufficient material condition of adequacy for
the definition of denotation the disquotational schema of which “The word
"Quine" denotes Quine” is an instance.

Of other semantical concepts besides truth, the notion of ontological
commitment is especially closely related to the concept of satisfaction, a
concept which was central in Tarski’s theory of truth and almost more fund-
damental in Tarski’s overall semantics theory than the concept of truth.
Today instead of satisfaction we usually speak equivalently of truth relative
to an assignment. An existentially quantified sentence $(\exists x)\phi$ is ontologically
committed to those entities which would satisfy the formula $\phi$ (when $x$ is
the only variable in $\phi$).\[82\]

It follows from the analogy between Church’s criterion of ontological
commitment and Convention T that Church’s criterion should be just as
convincing and indeed trivial as Convention T is usually thought to be. Of
course there have been many philosophers who have doubted the adequacy
of Tarski’s theory of truth, and these will no doubt remain similarly sceptical
about the adequacy of Quine’s theory of ontological commitment. However,
those that think that Tarski’s theory of truth is adequate should have no
reason to doubt the adequacy of the criterion of ontological commitment
as formulated by Church either. Similarly, problems about the interpreta-
tion and philosophical implications of Tarski’s theory will carry over to this
theory of ontological commitment.

\section{5.12 The Symbolization of ontological commitment}

The concept of ontological commitment is usually in the literature expressed
in natural language. However, if we are to develop a rigorous theory of
ontological commitment, we must make use of artificial languages. Since the
standard language used in philosophical logic is the language of predicate
logic, usually expanded in some fashion to contain non-logical constants, to
do so we should first ask how is the concept of ontological commitment to

\[82\] This must of course not be understood so that ontological commitment would be a
relation between a sentence and those entities which satisfy a formula. We can be ontolog-
ically committed to entities which do not exist and so though a sentence is ontologically
committed to entities which satisfy an embedded formula yet it may be that no entities
satisfy that formula. Because of this ontological commitment cannot be defined with
the aid of satisfaction at one blow, but we need a recursive or otherwise more complex
definition.
be symbolized in languages based on the language of predicate logic.

Since as we have seen the concept of ontological commitment is a semantical concept, analogous to the concept of truth, one might suspect that it requires a sharp distinction between language and metalanguage, and that neglecting such a distinction will lead to trouble. I will show that this is indeed the case; however, in order to show this, I will at first develop what may be called the naive theory of ontological commitment, in analogy to what are commonly called naive set theory and naive theory of truth (though in truth these are already rather sophisticated theories, sophisticated enough that they are hard to learn and hard to popularize, and sophisticated enough to lead to generate subtle contradictions, though not sophisticated enough to escape from them). I will therefore neglect at first the distinction between use and mention, in order to show that this will lead to paradoxes, and then later develop the theory of ontological commitment in a more sophisticated manner.

One natural first guess might be that ontological commitment should be symbolized as an ordinary predicate. If this were correct, then a sentence like "Sentence S carries ontological commitment to dogs." should be symbolized as \((\exists x)(D(x) \& C(S,x))\). However, a short reflection is enough to convince us that this is not an adequate symbolization. A sentence such as "There are dogs." carries ontological commitment to dogs, but it does not carry ontological commitment to any particular dog, as the proposed symbolization would do, for the sentence can be true even if any particular dog does not exist. We need therefore a more complicated symbolization. The second argument of the expression "ontological commitment" in natural language is typically a plural noun phrase, e. g. in the case of this example "dogs" or in a more complicated example "male brown dogs". Such phrases are often translated into predicates of predicates, such as \((\lambda P)(((\exists x)(D(x) \& P(x))))\) or even higher order constructs in rigorous formal systems of natural language semantics such as Montague semantics or Church’s own logic of sense and denotation. If we had to make use of this kind of treatment, then the theory of ontological commitment would become immensely difficult. However, such higher-order formalizations are typically only an intermediate phrase in symbolization; after symbolizing a sentence by means of such constructions, it is then simplified by means of such operations as \(\lambda\)-conversion into simpler forms. Using such symbolizations would be necessary only
if we were to seek for a way of mechanically formalizing all natural language sentences, as is typically done in Montague semantics; Montague’s compositional semantics is one promising way of producing such mechanical translations, though not the only one. While finding such mechanical translations is naturally desirable in the long run, it would be premature to do so in this stage of developing the theory of ontological commitment. I think it is enough to take the second argument to be either a predicate or an open formula (with the variable ranging over the entities of the kind to which the sentence carries ontological commitment). Open formulas can be converted into predicates by means of $\lambda$-conversion, so it suffices to take the second argument to be a predicate. The general form of statements of ontological commitment is then $C(x, s, \psi)$ or $C(s, P)$ where the former can be simplified into $C(s, (\lambda x)\psi))$. This can be read as ”The sentence $s$ carries ontological commitment to such entities $x$ that $\psi$.” Using this kind of symbolization, Convention C can now be stated in the following succinct form: $C((\exists x)\phi, (\lambda x)\phi)$. We can use $CE(\phi, (\lambda x)\phi)$ to symbolize ”Sentence $\phi$ carries explicit ontological commitment to such entities $x$ that $\phi$. In the case of explicit ontological commitment, Convention C can be expressed in a stronger form as the biconditional $(\forall P)(CE((\exists x)\phi, P) \equiv P = (\lambda x)\phi))$. Clearly this implies the weaker form $CE((\exists x)\phi, (\lambda x)\phi)$, as it can be derived from it by the reflexivity of identity, Universal instantiation, the definition of implication and Modus Ponens. Since $(\forall P)(CE((\exists x)\phi, P) \equiv P = (\lambda x)\phi))$ implies $CE((\exists x)\phi, (\lambda x)\phi) \equiv (\lambda x)\phi = (\lambda x)\phi)$ by universal instantiation, and this implies $(\lambda x)\phi = (\lambda x)\phi) \rightarrow CE((\exists x)\phi, (\lambda x)\phi)$ by the definition of equivalence and elimination of conjunction and since $(\lambda x)\phi = (\lambda x)\phi)$, we can derive $CE((\exists x)\phi, (\lambda x)\phi)$ by Modus Ponens.

While the concept of ontological commitment is intensional (as I will explain at greater length later) we can define with its aid an extensional concept, the extensionalization of ontological commitment. This extensional concept cannot replace the intensional concept, but will help us to apply extensional logic to the concept of ontological commitment in some situations. Let us say that an entity $a$ is among those entities to which a sentence carries ontological commitment if and only if $a$ is one of whatever entities the sentence is committed to. Let us use $CX(p, x)$ to symbolise that $x$ is among the entities to which the sentence $p$ carries ontological commitment.

If we are allowed to use besides a semantically closed language also
higher-order quantification (whether objectual or not), then this concept can
be defined so that $CX(p, x) \equiv_{def} (\exists P)(CE(p, P) \& P(x)) \& (\forall P)(CE(p, P) \to P(x))$, where $CE(p, P)$ means that the sentence $p$ carries explicit ontological commitment to entities $x$ such that $P(x)$. Thus e. g. if $p$ is the sentence "There are men.” and $x$ is a man, then any $x$ is among the entities to which $p$ carries ontological commitment iff $x$ is a man. Thus any man is among the entities to which $p$ carries ontological commitment, though we cannot say that $p$ would carry ontological commitment to him, since the sentence would be true even if he did not exist but some other man existed. To give a slightly different example, if $p$ is the sentence "There are unicorns.”, then $x$ is among the entities to which $p$ carries commitment if and only if $x$ is an unicorn. Thus since no unicorns exist no entity is among the entities to which $p$ carries commitment, but $p$ still carries commitment to some entities, namely to unicorns. It is now intuitively true that for any $p$ of the form $(\exists x)\phi$, $y$ is among the entities to which $\phi$ carries ontological commitment if $\phi(y/x)$, i. e. $CX((\exists x)\phi, y) \equiv \phi(y/x)$.

Clearly this intuitive truth is very close to the naive Convention C - as close as any extensional schema can be - and closely analogous to Convention T. However, in order to be more certain of this we can deduce it from Convention C and the inference rule of $\lambda$-conversion which is common in higher-order logics.

As the theorem to be proved, $CX((\exists x)\phi, y) \equiv \phi(y/x)$, is an equivalence, we can prove it by proving the implication in two directions, i. e. both $CX((\exists x)\phi, y) \to \phi(y/x)$ and $CX((\exists x)\phi, y) \leftarrow \phi(y/x)$. I will prove first $CX((\exists x)\phi, y) \equiv \phi(y/x)$, as this is a bit easier to prove. To prove an implication, I will suppose that its antecedent is true and prove on the basis of this assumption that its consequent is true, and then discharge the assumption by the introduction rule for implication. Let us then suppose that $CX((\exists x)\phi, y)$. By the definition of $CX$, replacing the definiendum by its definiens, $(\exists P)(CE((\exists x)\phi, P) \& P(x)) \& (\forall P)(CE((\exists x)\phi, P) \to P(x)))$. By the elimination of conjunction we can derive from this $(\forall P)(CE((\exists x)\phi, P)) \to P(x))$. By Universal Instantiation i. e. the elimination of the universal quantifier we can deduce from this $CE((\exists x)\phi, (\lambda y)(\phi(y/x)) \to (\lambda y)(\phi(y/x))(x)$ by replacing the predicate variable $P$ with the $\lambda$-abstract $(\lambda y)(\phi(y/x))$, which is a constant predicate term and so an appropriate term to substitute for it. Now $CE((\exists x)\phi, (\lambda y)(\phi(y/x)))$ is an instance of the weaker form

360
of Convention C (renaming systematically the bound variable \( y \)), which was one of our premises. Applying Modus Ponens i.e. the elimination of implication to this instance of Convention C and the result we arrived at so far (i.e. \( CE((\exists x)\phi, (\lambda y)\phi(y/x)) \rightarrow (\lambda y)(\phi(y/x))(x) \)), we can derive \((\lambda y)(\phi(y/x))(x)\). From this we can derive by \( \lambda \)-conversion i.e. the elimination of the \( \lambda \)-operator \( \phi(y/x) \). From the preceding sequence of stages of or proof, we can derive by the introduction of implication \( CX((\exists x)\phi, y) \rightarrow \phi(y/x) \). To prove the implication in the other direction, we start by assuming that \( \phi(y/x) \). From this we get by applying \( \lambda \)-conversion in the opposite direction \( (\lambda y)(\phi(y/x))(x) \). Now \( CE((\exists x)\phi, (\lambda y)\phi(y/x)) \) is an instance of the weaker form of Convention C (which we already proved to follow from the stronger form), so by applying the introduction of conjunction we get \( CE((\exists x)\phi, (\lambda y)\phi(y/x)) \& (\lambda y)(\phi(y/x))(x) \) and from this we derive \( (\exists P)(CE((\exists x)\phi, P) \& P(y)) \) by the introduction rule of the (higher order) existential quantifier and from this we get by the introduction of implication \( (\exists P)(CE((\exists x)\phi, P) \& P(y)) \leftrightarrow \phi(y/x) \). In order to prove that \( (\forall P)(CE((\exists x)\phi, P) \rightarrow P(x)) \) we have to appeal to the stronger form of Convention C, namely \( (\forall P)(CE((\exists x)\phi, P) \equiv P = (\lambda x)\phi) \). Let \( Q \) be an arbitrary property. We can derive by existential instantiation from the stronger form of Convention C that \( CE((\exists x)\phi, Q) \equiv Q = (\lambda x)\phi(y/x) \). We can derive from this by the definition of equivalence \( CE((\exists x)\phi, Q) \rightarrow Q = (\lambda x)\phi(y/x) \). Let us assume that \( CE((\exists x)\phi, Q) \). We can now derive by Modus Ponens \( Q = (\lambda x)\phi(y/x) \) and from this by the substitution of variables \( Q = (\lambda y)\phi(y/x) \). However, since as we already proved that \( (\lambda y)(\phi(y/x))(x) \), so by Leibniz’s Law \( Q(x) \). From this we derive by the introduction of implication \( CE((\exists x)\phi, Q) \rightarrow Q(x) \) (discharging the assumption \( CE((\exists x)\phi, Q) \)) and from this by the introduction of universal quantification since \( Q \) was an arbitrary property, \( (\forall P)(CE((\exists x)\phi, P) \rightarrow P(x)) \). From this we can finally derive by the introduction of implication (discharging the assumption that \( \phi(y/x)(x) \) \( (\forall P)(CE((\exists x)\phi, P) \rightarrow P(x)) \leftrightarrow \phi(y/x) \). Combining this with our earlier result that \( ((\exists P)(CE((\exists x)\phi, P) \& P(y)) \leftrightarrow \phi(y/x)) \) we arrive at \( ((\exists P)(CE((\exists x)\phi, P) \& P(x)) \& (\forall P)(CE((\exists x)\phi, P) \rightarrow P(x)) \leftrightarrow \phi(y/x) \) and replacing in the definition of the definition of \( CX \) the definiens with the definiendum, we finally arrive at \( CX((\exists x)\phi, y) \leftrightarrow \phi(y/x) \). Combining this with the earlier result that \( CX((\exists x)\phi, y) \rightarrow \phi(y/x) \) we get by the definition of equivalence \( CX((\exists x)\phi, y) \equiv \phi(y/x) \). Q.E.D.
5.12.1 A Paradox of Ontological Commitment

In his formulation Church carefully distinguishes between use and mention, between variables and their names and between open sentences and their names (even though his notation does not make the distinction as clear as it could be). This almost amounts to a distinction of an object language and a metalanguage, and can easily be generalised to such a distinction. Why does Church think that such distinctions are needed? Are such distinctions just pointless pedantry? No, Tarski made such distinctions in the theory of truth to make the theory of truth rigorous or formally adequate, and most importantly in order to prevent semantical paradoxes such as the Liar paradox. It seems to me that Church may have known or suspected that similar paradoxes threaten us in the case of the notion of ontological commitment.

If this was what Church was concerned about Church was right; without distinctions between object language and metalanguage we can easily be led into contradiction in the case of ontological commitment exactly as in the case of truth. We can formulate in our naive theory of ontological commitment a paradox similar to the Paradox of the Liar; I will call it the Paradox of Denied Ontological Commitment. Consider the following example sentence:

Example 11 There are entities to which this sentence carries no ontological commitment.

This sentence\textsuperscript{83} is very similar to the Liar sentence "This sentence is not true." When we ask ourselves whether the Liar sentence is true or not, it seems intuitively that it must be true just in case it is not true. From this, however, it follows that if the Law of the Excluded Middle is true, the sentence must be both true and not true, which is a contradiction. Similarly when we ask to which entities the sentence above carries ontological commitment, it seems intuitively - to put this very crudely and inaccurately - that this assertion would carry ontological commitment to any entities just in case it does not carry ontological commitment to them.

In order to derive an explicit contradiction in a rigorous way we have to use a slightly more complicated route, making use not only of the concept

\textsuperscript{83}We could of course also use the concept of pragmatical ontological commitment in deriving a paradox; the sentence "There are entities to which I am not ontologically committed." would do as well, except that it is not quite as simple semantically (even though its expression is shorter).
of ontological commitment but of a more complicated concept that can be
defined with the aid of it, namely the concept of an entity being among the
entities a sentence is ontologically committed to. We have already defined
this concept, and proved a theorem regarding it; we can easily see with the
aid of this theorem that applying this concept to Example 11 leads to an
explicit contradiction. Thus for example any man is among the entities to
which the sentence Example 11 carries ontological commitment just in case
he is not among them. If we assume that any man must either be among the
entities to which the sentence carries commitment or not be among them,
which is just an instance of the Law of the Excluded Middle, we can easily
derive an explicit contraction; any man both is and is not among the entities
to which the sentence carries commitment.

The parallel to truth is perhaps not wholly complete in this respect
since it may be that a contradiction cannot be derived from the criterion
of ontological commitment in every language which is semantically closed
in the sense that we can speak in it of the ontological commitments its
own sentences carry and which contains the standard propositional logical
connectives, but only in a sufficiently rich semantically closed language, a
language with higher-order quantifiers or plural quantifiers (which perhaps
need not be objectual, so that their addition may not have to increase the
ontological commitments made in the language). However, since very many
such languages do not lead to any contradiction without the addition of the
concept of ontological commitment\footnote{Of course a language which has higher-order quantifiers together with a naive, too
strong comprehension principle leads to contradiction even without the addition of any
semantical concept such as the concept of ontological commitment, but my point is that
even a language without any comprehension principle or any similar controversial prin-
ciples which does not harbour any contradictions will lead to contradictions when the
concept of the ontological commitment of its own sentences is added to it together with
intuitively obvious axioms concerning it.}, and do lead to contradiction when
the concept is added along with its criterion, the parallel is in any case very
far-going.

I will now show how a contradiction can be explicitly derived in a lan-
guage which contains the extensionalization of the concept of ontological
commitment. As we have already proved, $CX((\exists x)\phi, y) \equiv \phi(y/x)$. Now let
c be the sentence such that $c = (\exists x)(\neg CX(c, x))$. Clearly such sentences
could exist in a semantically closed language in which more familiar para-
doxes like the Liar paradox could be formulated as soon as the concept of

363
truth was introduced to it. It is quite similar to the Liar sentence \( l \), which is such that \( l = \neg \text{True}(l) \), being just a slightly more complicated instance of an expression talking about its semantical values. Now it follows from what we have said that \( CX((\exists x)(\neg CX(c, x)), y) \equiv \neg CX(c, y) \), i.e. since \( c = (\exists x)(\neg CX(c, x)) \), \( CX(c, y) \equiv \neg CX(c, y) \). This final sentence is contradictory according to classical two-valued gapless propositional logic\(^{85}\).

It would be disingenuous to conclude from these paradoxes that there is something wrong with the concept of ontological commitment, unless one is willing to say that there is also something wrong with the notion of truth, which is surely absurd. Obviously most if not all ways suggested to prevent the Paradox of the Liar also suffice to prevent this paradox and help to make the theory of ontological commitment formally correct. The obvious solution is of course to demand that we could not speak in any language of the ontological commitment of its own sentences even as Tarski demands that we cannot speak in any language of the truth of its own sentences\(^{86}\). Church seems to strongly suggest this solution, though he does not explicitly argue for it. We must then make a distinction between object language and metalanguage, or more generally between languages and their metalanguages. We will then have a first-level language which cannot speak about any ontological commitments though it has such, a second-level language in which we can speak about the ontological commitments of sentences in the first-level language, a third-level language in which we can speak about the ontological commitments of sentences of the second-level language (and also the second-order language, if we make the hierarchy of languages cumulative, as we should in the interest of expressive power), and so on.

However, though this is the most obvious solution it is not the only one and is not necessarily the best. Alternatively we might choose a solution like Kripke’s in [Kri75] or a solution like the revision theory of truth. In these solutions the predicate of ontological commitment, like that of truth, would

\(^{85}\)In order to derive an explicit contradiction we must assume that there exists some entity \( y \). This assumption is trivially true in standard predicate logic, but not in free logic. However, it is obviously a true assumption; one can hardly seriously claim that the contradiction could be avoided by assuming that nothing exists. Besides, \( y \) I might be identical with the sentence \( c \) itself; a semantically closed language must be capable of speaking about its own sentences, and therefore a theory formulated in it must assume that sentences exist.

\(^{86}\)This seems to be contrary to the claim of Devitt in [Dev84, §4.6 pages 50-53] that we can establish ontological commitments without moving into a metalanguage discussion, though the word "establish" that Devitt uses is rather ambiguous.
be defined in stages, which correspond to the levels of the Tarskian hierarchy of languages but are all stages in the definition of a single language. In a Kripkean solution the Law of Excluded Middle is denied, and the assertion Example 11 would turn out to be without truth-value. I will for the sake of simplicity use the more familiar Tarskian approach in the rest of this dissertation, but in the end I suspect some more complicated approach such as that of Kripke will be better, as it captures better the way truth-predicates are used in natural languages. By this I do not mean just that it describes the way they function better (though it does that) for describing the way natural languages actually work is not the only or even the primary task of semantics, but also that it provides a better replacement for the inconsistent concept of truth used in natural languages (I mean a better replacement in the sense that we can do with it more of the jobs that inconsistent concept tried to do while retaining consistency). Nevertheless, even a Kripkean approach cannot entirely do away with the use of metalanguages, since such things as that some sentences of the object language do not possess a truth-value, cannot be said in that language itself, so it remains firmly based upon Tarskian foundations.

5.13  A Recursive Characterisation of Explicit Ontological Commitment

It is often stated (most prominently by Horwich) to be the main idea of deflationist theories of truth that the Convention T should be taken as an axiom schema of the theory of truth, indeed as the only axiom schema of such a theory, rather than just as a material condition of adequacy for the theory of truth. If as I argue the Convention C is analogous to Convention T, then a deflationary theory of ontological commitment should likewise take the Convention C as an axiom schema rather than just as a criterion of adequacy. This would result in a very simple theory of ontological commitment, as Horwich’s theory results in a very simple theory of truth. However, just as it has been often questioned whether a theory of truth as simple as Horwich’s is adequate, it can be questioned whether such a theory of ontological commitment would be adequate.

We have seen that a purely disquotational theory of truth would hold that a person can apply the truth predicate only to sentences he under-
stands. Analogously a purely disquotational theory of ontological commitment would say that a person can apply a predicate of ontological commitment only to sentences he himself understands, while a more inflationary theory of ontological commitment would say that a person can speak about the ontological commitment of sentences which he cannot understand, indeed, even of the ontological commitment of sentences he cannot understand. However, as we have seen even Field, who introduced the pure disquotational theory of truth has realized that it is not adequate, so there is no point in trying to construct a purely disquotational theory of ontological commitment.

It is in fact easy to see how to develop a deflationary theory of ontological commitment. As I will show, the concept of ontological commitment can be defined using the concept of reference (or designation). However, there exist many deflationary concepts of reference; most philosophers who are deflationists about truth are also deflationists about reference, though deflationism about truth does not imply deflationism about reference (though the converse relationship probably holds). For example, Robert Brandom has in [Bra84] developed an anaphoric theory of reference which would make reference an intra-linguistic relationship, and Paul Horwich has developed a less radical deflationary theory of reference. If the concept of reference used in my definitions were to be reinterpreted in a deflationary way then my theory of ontological commitment would be transformed into a deflationary theory.

However, this would trivialize the concept of ontological commitment and of ontology itself completely (at least if the deflationary theory in question were radical like Brandom’s). In fact, the result of this kind of procedure would be similar to Quine’s thesis of ontological relativity, which is not surprising since Quine is a deflationist. However, it is surely not necessary to interpret the concept of reference in a deflationary way and I do not want to do so, since I want to develop ontological theories that are not completely trivial.

Tarski’s Convention T was not, however, the culmination of Tarski’s theory of truth. It was not itself a definition of truth in the sense in which Tarski was seeking for a definition of truth but only a preliminary criterion for the material adequacy of a definition of truth. The Convention T was according to him something to be explained, not indeed causally (for it is not
a historical event such as could be explained causally, but a timeless general rule) and not directly naturalistically (for even if a naturalistic explanation of it is in principle possible, we are far from being able to give such an explanation today) but in terms of more fundamental semantical principles. This is because the notions that occur in the formulation of Convention T are not to be taken as primitive notions either. Tarski did not take the notion that ‘X’ is the name of ‘p’, which he used in his formulation of Convention T, as a primitive - not in the case where ‘p’ is a sentence - but characterised it further by using other semantical notions such as satisfaction and the denotation of simpler expressions. It was only in the case of the denotation of simple individual and predicate constants that Tarski resorted to disjunctive definition, which if Field is right amounted to leaving these semantical notions primitive. In the case of complex expressions such as sentences Tarski was able to give a non-disjunctive, non-trivial definition for all semantical concepts applying to them which he used.\footnote{This is at least in part why Tarski’s theory may not be deflationist in any strong sense of the word. Deflationists typically think that only an implicit definition of truth is possible, but Tarski’s theory gives an explicit definition of truth (relative to particular languages).}

If the analogy between truth and ontological commitment is valid, then the Convention C or C-schema can also be only a condition for the material adequacy of a theory of ontological commitment, not a complete theory of ontological commitment in itself. Carrying the analogy further, an explicit, non-disjunctive definition ought to be possible also in the case of ontological commitment. We must therefore proceed further if we are to find a fully satisfactory theory of ontological commitment. We must explain Church’s notion of the name of a propositional form further, since we can take the name-relation i. e. the relation of denotation as primitive only in the case of simple individual and predicate constants, not in the case of any complex formulas (or complex terms either, if the language has them). However, the way ahead is not very difficult, as we have just to follow the obvious analogies with other semantical concepts like truth and denotation. If a basically Tarskian approach to the problem of truth is correct, as I have already argued, then because of the analogies between the notions of truth and ontological commitment, the notion of ontological commitment should be defined like the notion of truth recursively by means of denotation, and Convention C should be proved on the basis of these definitions rather than
just taken as an axiom schema.

One cannot indeed very naturally speak about a correspondence theory of ontological commitment, for since we can be ontologically committed to entities which do not and even cannot exist, ontological commitment cannot be a matter of a straightforward correspondence between linguistic expressions and entities in the world. Nevertheless, one can ask whether Quine’s or Church’s theory of ontological commitment is a deflationary or an inflationary (or substantive or substantivalist) theory of ontological commitment, and more importantly, whether the theory of ontological commitment should be deflationary or inflationary (substantive or substantivalist).

So the definition of ontological commitment must be recursive like the definition of truth. We must also initially relativise the notion of ontological commitment to sequences of entities as Tarski did or equivalently to assignments of variables as is more commonly done today. I will use the metavariables $g, h, \ldots$ to range over assignments, i.e. functions which associate entities with variables of the object language\(^{88}\). Let us say that for any term $t_i$ (including both singular terms and predicate terms) $\|t_i\|_g$ is the semantic value of $t_i$ relative to assignment $g$, i.e. $\|t_i\|_g = \alpha$ iff either $t_i$ is a variable and $g(t_i) = \alpha$ or $t_i$ is a constant and $t_i$ is the name of $\alpha$.

A sentence carries commitment to entities $x$ such that $M$ if and only if it carries commitment to entities $x$ such that $M$ relative to all assignments $g$. While Tarski’s Convention T applied directly to all sentences, Church’s Convention C applies directly only to existentially quantified sentences. Thus while Tarski’s recursive definition of truth ranges over all forms of formulae from the simpler to the more complex, our definition of ontological commitment must range (at least at first) only over all existentially quantified formulae. However, since existential quantification can be applied to any formula, this difference is smaller than might at first appear. This implies that the recursion must range over all formulae occurring in the scope of an

---

\(^{88}\)One might object that the common use of assignments involves stronger ontological commitments than the use of sequences, since it involves commitments to sets of sequences (as an assignment is such a set) and not just to sequences, and that we ought to avoid such commitments if they are not necessary. However, a Tarskian theory in any case carries commitments to sets of sequences as the extensions of polyadic predicates, and therefore we cannot avoid such commitments anyway if we are to develop a Tarskian theory further rather than radically revise it. Therefore we can just as well embrace them already by using assignments, and this is useful since this avoids some of the complications of Tarski’s original formulation of his theory by means of sequences (such as the use of infinite sequences).
existential quantifier, from the simpler to the more complex. Obviously we must start from atomic formulae occurring in the scope of the existential quantifier.

Let $\|t_i\|$ stand for the designation (aka denotation aka reference aka extension) of the term $t_i$ relative to assignment $g$.

**Preliminary Definition 1** Sentence $(\exists x)R(t_1, \ldots, t_n)$ where the quantifier $(\exists x)$ is objectual carries ontological commitment to such entities $\alpha$ that $\langle t_1\|_{g(\alpha/x)}, \ldots, t_n\|_{g(\alpha/x)} \rangle \in \|R\|_{g(y/x)}$ relative to assignment $g$.

We can replace the implication here with an equivalence if we restrict ourselves to explicit ontological commitment. Let $R$ be the name of the relation $Q$ or, to say this in terms not committing us to realism, \('R'\) be the name of \('Q'\).

**Metadefinition 3** A sentence $\phi$ carries explicit ontological commitment to entities $\alpha$ such that $Q(\beta_1, \ldots, \beta_n)$ where $\alpha = \beta_i$ for some $1 \leq i \leq n$ relative to assignment $g$ if and only if it is identical with such a sentence $(\exists x)R(t_1, \ldots, t_n)$ that the quantifier $(\exists x)$ is objectual and unrestricted and $\|R\|_g = Q$ and $\beta_1 = \|t_1\|_g, \ldots, \|x\| = \alpha, \ldots$ and $\beta_n = \|t_n\|_g$.

It must be stressed that a sentence $(\exists x)R(t_1, \ldots, x, \ldots, t_n)$ can carry explicit ontological commitment to such entities $y$ that $(\|R\|_g(t_1\|_g, \ldots, y, \ldots, \|t_n\|_g))$ according to this definition even if the sentence is false and there are no such entities, i. e. even if for no $y$ $(t_1\|_g, \ldots, \|y\| \in \|R\|_g)$. Furthermore, it can carry that ontological commitment even if there could not any such entities, i. e. even if $(\exists x)R(t_1, \ldots, x, \ldots, t_n)$ is not true even with respect to any model or any possible world. Thus, explicit ontological commitment is a hyperintensional notion in the sense in which the word is used by Creswell in [Cre75], even though we have not had to use any intensional notions in characterizing it.\(^{89}\)

\(^{89}\)It might be possible to give a non-recursive characterization of explicit ontological commitment by using notions developed in the tradition of hyperintensional logic. Two sentences have the same explicit ontological commitments only if they are intensionally isomorphic in the sense of the word Carnap developed in [Car47]. The recursive semantic rules characterizing explicit ontological commitment work because they create an isomorphism very similar to intensional isomorphism between the expressions in the object language and the metalanguage. More exactly, the sentence "The sentence $S$ carries explicit ontological commitments to such entities $x$ that $M'$ is true if and only if the sentences $S$ and $(\exists y)M$ are intensionally isomorphic. However, there is a reason why the recursive def-
The next case is that where a negative formula is in the scope of the existential quantifier.

**Metadeﬁnition 4** A sentence \( \phi \) carries explicit ontological commitment to entities \( \alpha \) such that it is not the case that \( M \) relative to assignment \( g \) if and only if it is (identical with) such a sentence \((\exists x)\overline{\psi}\) that (the quantifier \((\exists x)\) is objectual and unrestricted) and the sentence \((\exists x)\psi\) carries ontological commitment to entities \( \alpha \) such that \( M \) relative to assignment \( g \).

For example, let us assume that \( \psi \) is \( P(x) \) and \( \|P\|_g \) denotes the set of material beings or the property of being material. In this case \( \phi \) is the sentence \((\exists x)\overline{P(x)}\) and carries ontological commitment to immaterial beings (whether such beings exist according to the metatheory or not).

Adding explicitly that the quantifier is objectual and unrestricted may not be necessary in this case, or in any of the following cases of complex formulas, since we are supposing that the quantifier applied to the negative formula is the same as the one applied to the negated one, and since the later is objectual and unrestricted according to the other rules (by the recursion which starts from the previous rule) it follows that the former must also be objectual and unrestricted.

In the next case a conjunctive formula is in the scope of the existential quantifier.

**Metadeﬁnition 5** Sentence \( \phi \) carries explicit ontological commitment to entities \( \alpha \) such that \( M \) and \( N \) relative to assignment \( g \) if and only if it is (identical with) such a sentence \((\exists x)(\psi \land \chi)\) that the quantifier \((\exists x)\) is objectual and unrestricted and the sentence \((\exists x)\psi\) carries ontological commitment to entities \( \alpha \) such that \( M \) relative to \( g \) and sentence \((\exists x)\chi\) carries ontological commitment to entities \( \alpha \) such that \( N \) relative to \( g \).

For example, the sentence "There are flying horses." carries commitment to flying horses. Let us assume that \( \psi \) is the formula \( P(x) \) and \( \chi \) is the...
formula $Q(x)$. Let us further assume that $\|P\|_g$ stands for the set of entities capable of flying or the property of being capable of flying and $\|Q\|_g$ for the set of horses or the property of being a horse. In this case $\phi$ is the sentence $(\exists x)(P(x) \land Q(x))$ and carries ontological commitment to flying horses (even if they do not exist according to the metatheory).

It must be noted that the conjunctive formula may be even contradictory. A formula such as $(\exists x)(P(x) \land \neg P(x))$ carries commitment to logically impossible entities (e.g., if $P$ stands for men or humanity to beings which are men yet are not men). Despite this, however, it does not carry commitment to entities satisfying an arbitrary contradictory formula. This makes even clearer the hyperintensional character of the theory of explicit ontological commitment, and actually makes the theory paraconsistent.

Next follow sentences where a disjunction is in the scope of the existential quantifier.

**Metadefinition 6** Sentence $\phi$ carries explicit ontological commitment to entities $\alpha$ such that $M$ or $N$ relative to assignment $g$ if and only if it is identical with such a sentence $(\exists x)(\psi \lor \chi)$ that the quantifier $(\exists x)$ is objectual and unrestricted and the sentence $(\exists x)\psi$ carries ontological commitment to entities $\alpha$ such that $M$ relative to assignment $g$ and sentence $(\exists x)\chi$ carries ontological commitment to entities $\alpha$ such that $N$ relative to assignment $g$.

For example, the sentence "There are trees or mountains." carries commitment to entities which are trees or mountains, i.e. it carries commitment to trees or mountains. If $\psi$ is again $P(x)$ and $\chi$ is $Q(x)$ and $\|P\|_g$ denotes the property of being a tree and $\|Q\|_g$ denotes the property of being a mountain, then $(\exists x)(P(x) \lor Q(x))$ carries ontological commitment to trees or mountains.

While sentences with explicit ontological commitment are all quantified, some of them may yet have more than one quantifier. Because of this we must consider quantified sentences separately. In the following $g(\alpha/x)$ is the assignment that is otherwise the same as $g$ except that it associates the variable $x$ of the object language with the entity $\alpha$.

**Metadefinition 7** Sentence $\phi$ carries ontological commitment to entities $\alpha$ such that for some entity $\beta$ $M(\beta/\alpha)$ relative to assignment $g$ if and only if it is identical with such a sentence $(\exists x)(\exists y)\psi$ that the quantifier $(\exists y)$ is
objectual and unrestricted and for all entities $\beta$ the sentence $(\exists x)\psi$ carries ontological commitment to entities $\alpha$ such that $M$ relative to assignment $g(\beta/\alpha)$.

An example may make this clearer. If $\psi$ is $M(x) \land L(x, y)$, then $\phi$ is $(\exists x)(\exists y)(L(x, y))$. If $\|L\|_g$ is the relation of loving, then $\phi$ carries ontological commitment to men who love some woman.

### 5.14 From Explicit to Implicit Ontological Commitment

Though we have thus gone through all the complex sentences that can be formed in standard logical languages, we have not yet completed our task. This is because Church’s Convention C itself is incomplete. There is one way in which Church’s formulation of the criterion of ontological commitment as it stands is obviously incomplete and indeed is not even as good as Tarski’s Convention T, though it turns out that this defect can be rather easily corrected - at least in part. Church’s criterion gives only sufficient, not necessary conditions for ontological commitment. Therefore it does not enable us to say what entities an assertion which is not of the form $(\exists x)M$ is ontologically committed to. Nor does it tell exhaustively even what an assertion which is of this form is ontologically committed to. Even if an assertion carries commitment to entities $x$ such that $M$, it does not follow that it could not carry commitment to other entities, even entities which are not such that $M$.

This is easy to see with the aid of an analogy; if a man likes strawberries it does not follow that he does not like anything else; he might still like also blueberries, even though blueberries are not strawberries. It

---

90 Chihara apparently thinks that this would follow; he says in [Chi74, page 77] that Church’s criterion does not allow us to say that ‘$(\exists x)(x \text{ is a set})$’ carries ontological commitment to an entity $x$ such that $x$ is an abstract entity. However, Church’s criterion does allow us to say that, though it does not force us to say that. Church’s formulation just leaves open the question whether the sentence in question carries commitment to abstract entities, which shows that Church’s formulation is incomplete, but not that there would be a positive error in it.

91 The point of the analogy is that the phrases “likes” and “is ontologically committed to” in English are similar both syntactically and semantically. Syntactically both can take as complements a full noun phrase, but also any phrase projected by a plural or singular noun. Semantically both are objectual attitude verbs which create an intensional context, in which coextensive terms are not substitutable. Even if the first star that rises in the morning is the same as the last star that sets in the evening, one can like the first star
might of course be that Church intended to say that the assertion of $(\exists x)M$ carries ontological commitment to entities $x$ such that $M$ and to no other entities. However, this would not be a plausible position, and charity precludes interpreting Church as maintaining it, since there is no evidence he did so. Rather, Church just left open the question of what are the necessary conditions of ontological commitment.

Intuitively if a statement implies or presupposes another, its ontological commitments should include those of the one it implies. E. g. intuitively we would surely say that the statement "There are lions and there are baboons." would carry commitment to lions, just like the weaker statement "There are lions." since it implies this weaker statement. Similarly the statement "The king of France is bald." should carry commitment to a king of France since it presupposes that there is a king of France. However, assertions that imply or presuppose the assertion $(\exists x)M$ are not themselves of that form, so this does not follow from Church’s criterion. Therefore we must reformulate the criterion of ontological commitment so that this follows from it.

It is not hard to see in what kind of way we should modify Church’s formulation of his criterion so that it would give conditions which are not only sufficient but also necessary. The way in which this can be done can be found already by examining the connection between the commitments of assertions and theories which I have suggested, that the commitments of an assertion, i. e. a statement, are those of its logical closure. It follows from this together with Church’s criterion that even if a statement is not of the form $(\exists x)M$, yet if it implies such a statement, it also carries ontological commitment to entities $x$ such that $M$. Implying such a statement is a stronger sufficient condition of commitment to entities $x$ such that $M$ than the one Church proposed. I suggest it is also a necessary condition for such commitment. Therefore I suggest the following rough formulations of ontological commitment:

**Preliminary Definition 2** An assertion $S$ which is not of the form $(\exists x)M$ where the quantifier $(\exists x)$ is objectual carries ontological commitment to entities $x$ such that $M$ that rises in the morning but not like the last star that sets in the evening and similarly one can be ontologically committed to the first star that rises in the morning without being ontologically committed to the last star that sets in the evening. Nevertheless, the analogy must not be pressed too far; there are also significant differences. For example while liking is the mental state of a single person, ontological commitment is based on social conventions.
tities \( \alpha \) such that \( M \) if and only if it implies such an assertion \( N \) that \( N \) carries explicit ontological commitment to entities such that \( M \).

**Preliminary Definition 3** A theory \( S \) carries ontological commitment to entities \( \alpha \) such that \( M \) if and only if some sentence that is contained in or implied by it carries explicit ontological commitment to entities \( x \) such that \( M \).

These formulations are yet very rough for one crucial reason. I have left it open what concept of implication is used here. Obviously material truth-functional implication is not enough, since if it were any two true statements would have the same ontological commitments, which is utterly absurd. However, logical implication at least is enough for ontological commitment. For example, the sentence \((\exists x)(P(x) \& (\exists x)Q(x))\) is clearly committed to entities \( x \) such that \( P(x) \) as it implies the sentence \((\exists x)(P(x))\) and is also committed to entities \( x \) such that \( Q(x) \) as it implies the sentence \((\exists x)(Q(x))\).

To take an even more concrete example, the statement "There are men and lions." clearly carries commitment to men since it implies the statement "There are men" and it also carries commitment to lions since it implies the statement "There are lions".

In fact if we only want to capture as many as possible of Quine’s own intentions, then we probably should stop at logical implication. This also yields a concept of ontological commitment that is not wholly trivial, but actually has some uses. Thus we would say that an assertion \( S \) carries ontological commitment to entities \( x \) such that \( M \) if and only if \( S \) logically implies \((\exists x)M\), i.e. \( S \vdash (\exists x)M \). Let us call this rather narrow concept of ontological commitment implicit formal ontological commitment.

**Preliminary Definition 4** An assertion \( S \) carries implicit formal ontological commitment to entities \( x \) such that \( M \) if and only if \( S \models (\exists x)M \) where the quantifier \((\exists x)\) is objectual and unrestricted.

Making this fully precise involves introducing model-theoretic machinery.

**Preliminary Definition 5** An assertion \( \phi \) carries implicit formal ontological commitment to entities \( \alpha \) such that \( M \) relative to assignment \( g \) if and only if there is such a sentence \( \psi \) that \( \psi \) carries explicit ontological commitment to entities \( \alpha \) such that \( M \) relative to \( g \) and for all models \( \mathcal{M} \) and
assignments \( h \phi \) is true with respect to \( M \) and \( h \) only if \( \psi \) is true with respect to \( M \) and \( h \) (i.e. \( \models_{M,h} \phi \rightarrow \models_{M,h} (\exists x) \psi \)).

The conception of ontological commitment can also be related to syntactical concepts. It would follow from this definition that an assertion \( \phi \) formulated in a language which has a complete logic carries implicit ontological commitment to entities \( x \) such that \( M \) relative to assignment \( g \) if and only if there is such a sentence \( \psi \) that \( \psi \) follows from \( \phi \) according to the inference rules of that logic, i.e. \( \phi \vdash \psi \) and \( \phi \) carries explicit ontological commitment to entities \( x \) such that \( M \) relative to \( g \), i.e. \( \psi \) is of the form \((\exists x)M\). If the language in which the assertion is formulated does not have a complete logic, but has an incomplete logic (i.e. a logic whose theorems are all valid but which cannot prove some valid sentences), then an assertion formulated in that language carries implicit ontological commitment to entities \( x \) such that \( M \) relative to assignment \( g \) if (but not only if) there is such a sentence \( \psi \) that \( \psi \) follows from \( \phi \) according to the inference rules of that logic, i.e. \( \phi \vdash \psi \) and \( \psi \) carries explicit ontological commitment to entities \( x \) such that \( M \) relative to \( g \), i.e. \( \psi \) is of the form \((\exists x)M\).

There is one additional complication. While the above definition is already precise enough for most purposes it may not be wholly adequate. We must also take into consideration the possibility that the language whose ontological commitments we are examining may not have sufficient expressive resources to express explicitly the implicit commitments of statements formulable in it, so that it will have to be extended to a richer language to examine the ontological implications of its statements. This final alteration may be controversial, since it increases greatly the number of implicit formal ontological commitments a sentence can have. This alteration reflects the intuition that the implicit formal ontological commitments of a sentence follow solely from the interpretation of (the constituents of) that sentence and general logical rules, and it is irrelevant to them how rich is the expressive power of the rest of the total language of which the sentence is a part. Therefore linguistic holists might not like this change (so Quine himself might not have liked it), but it will seem intuitive to those whose semantical intuitions are more atomistic, as the compositional (or more generally recursive or inductive) intuitions at the basis of Tarski’s theory of truth arguably are. Also it may help in comparing the ontological commitments of incommensurable theories.
In order to avoid semantical paradoxes the languages into which the language whose ontological commitments we are examining can be expanded must be limited into those that are of a lower level in the Tarskian hierarchy than the metalanguage in which we are examining the ontological commitments of the first object language. This has, however, the somewhat paradoxical and unwelcome consequence that the ontological commitments of a language would be relative to a metalanguage; this relativity is far weaker than the kinds of ontological relativity that many philosophers such as the late Quine or even Chateaubriand have argued for, as it is a relativity only with respect to a metalanguage but not with respect to any metatheory, but still in my view unwelcome. Nevertheless, this kind of minor relativity seems inevitable given a standard solution to the semantical paradoxes. A better solution of the paradoxes, if available, might be capable of removing this kind of relativity; however, I will have to accept it for now. Let us assume that the metalanguage is of level $n$; then we can say the following:

**Metadefinition 8** An assertion $\phi$ formulated in a language $L$ carries implicit formal ontological commitment to entities $\alpha$ such that $M$ relative to assignment $g$ if and only if there is such a sentence $\psi$ in a language $L'$ extending $L$ (i.e. $L \subseteq L'$) that $L'$ is of a level $m < n$ and $\psi$ carries explicit ontological commitment to entities $\alpha$ such that $M$ relative to $g$ and for all models $M = \langle I, D \rangle$ where the interpretation function $I$ is defined on $L'$ (so that all the primitive expressions of $L'$ belong to its domain) and assignments $h \phi$ is true with respect to $M$ and $h$ only if $\psi$ is true with respect to $M$ and $h$ (i.e. $\models_{M,h} \phi \rightarrow \models_{M,h} (\exists \alpha) \psi$).

This definition of ontological commitment has many significant and perhaps startling implications. It can be used to show that a theory of modal logic, that only uses modal operators and in which quantification over possible worlds is not possible, will yet carry ontological commitments to possible worlds and possible individuals, if the semantics given to the modal operators is the usual possible worlds semantics. This follows from the fact that the language can be extended to one in which we can quantify over possible worlds, and all such expansions make true such sentences as $\Diamond(p) \rightarrow (\exists w)(In(w,p))$, where $In(w,p)$ means that $p$ is true in $w$. This does not necessitate an ascent in the Tarskian hierarchy of languages, since such an expansion does not make it possible to speak of the sentences of the
original language.

If we use this more complicated definition, the connection between semantics and syntax stated above must be slightly modified. It follows from this definition that even an assertion formulated in a language which has a complete logic carries implicit ontological commitment to entities \( x \) such that M relative to assignment \( g \) if (but not only if) there is such a sentence \( S' \) that \( S' \) follows from \( S \) according to the inference rules of that logic, i.e. \( S \vdash S' \) and carries explicit ontological commitment to entities \( x \) such that M relative to \( g \), i.e. \( S' \) is of the form \((\exists x)M\).

The definition of implicit formal ontological commitment may seem similar to that of logical entailment. However, we must not lose sight of the fact that though the definition of implicit ontological commitment makes use of the notion of logical implication, implicit ontological commitment is yet a material, not a formal and hence not a (in the narrow sense) logical notion; in this it is just like the notion of material truth that Tarski defined in his original theory of truth. It is essential for the definition that we make use in the metalanguage of the constant M, which is typically a non-logical constant. In the schema we have formulated above this constant is indeed quantified over, so the schema seems to be in a sense formal; however, the specific instances of the schema are definitely material truths. The schematic constant M is in the instances of the schema replaced by such material constants as "is a man" or "is red", etc.

5.15 Ontological Commitment in Natural Languages

The formulation of the criterion for implicit formal ontological commitment applies directly only to assertions and theories expressed in a language of predicate logic. Therefore it does not apply directly to all assertions made in scientific theories either, as most scientific theories, though they are expressed with the aid of a lot of mathematical symbolism, are not yet expressed solely in the language of predicate logic, and natural language expressions usually play some role in their expression. The formulation of the criterion for implicit material commitment, on the other hand, may be directly applicable to natural languages, but discovering how it is applicable is not very straightforward, but presupposes that a semantical theory has been developed of them. How these criteria are to be applied in detail to
assertions made in natural language (or vernacular, as it is sometimes called in the literature) and to scientific theories is a question to which I now turn.

Quine himself was in many writings sceptical about the very possibility of applying the concept of ontological commitment to natural languages in any exact way, though he applied it to them in the informal examples he gave. This skepticism is similar to the scepticism expressed by Tarski in \[\text{Tar83a}\] and \[\text{Tar44}\] about applying the concept of truth to natural languages in any exact way. However, after Tarski many other philosophers and linguists such as Karl Popper, Richard Montague, Donald Davidson\(^92\), Jaakko Hintikka etc. have quite successfully applied Tarski’s theory of truth to natural languages. As we have seen that there are deep-going parallelisms between the theory of truth and the theory of ontological commitment, we have therefore good reasons to except that Quine’s and Church’s theory of ontological commitment can be similarly generalized and applied to natural languages.

It is important to notice that natural language is not the same as ordinary language. Many philosophers who talk about ontological commitments in the vernacular, such as Jody Azzouni, seem to confuse these two concepts. Natural language comprehends also the technical jargon of scientists and other scholars, including the jargon of philosophers as one specific case of it. The very phrase “ontological commitment” is not part of ordinary language; the man in the street would usually have no idea what it means. It is, however, an English expression and therefore a part of a natural language. Ontologists and metaphysicians do not use ordinary language, though they generally use natural language. Of course, older metaphysicians and ontologists like Aristotle or Wolff who invented the very sciences did not yet have any artificial language to use. However, the language of Aristotle was probably very different from the colloquial Greek of his time and Wolff of course wrote in a language, Latin, that was dead as an ordinary language when he used it. They wrote in jargons. Jargons, of course, develop out of ordinary languages and have many similarities to them. However, they have also many differences, and may be as far from ordinary languages as regi-

\(^92\)Donald Davidson held that Tarski’s theory was applicable to natural languages both when he thought that it was a correspondence theory and when he no longer thought so. The question whether and how far Tarski’s theory can be applied to natural languages does indeed appear to be logically independent of the question whether it is a correspondence theory or not.
mented languages of symbolic logic are from them. Azzouni argues that no specific locutions carry ontological commitment in ordinary language; even if this argument were correct, it would not imply that no specific locutions carry ontological commitment in philosophical jargon, and hence it would not follow that there are no criteria of ontological commitment in natural language. Azzouni claims in [Azz04, page 117] that ontological indicators involve inextricably pragmatic indicators such as rhetorical stress. However, such pragmatic factors surely do not suffice us to express anything with the precision any science or scholarly discipline requires, and therefore cannot express ontological commitments with the precision a discipline requires either. However, philosophers often make ontological claims in articles and books, where such pragmatic indicators as stress are not usually to be had; philosophers may make use of such pragmatic indicators when lecturing, but they had better be able to dispense with them if they are ever to publish anything. Azzouni claims that ordinary speakers make - both consciously and tacitly - genuine ontological claims. I am not sure if this is correct, especially when it comes to conscious ontological claims, though I do not want to deny it. However, it is obvious that philosophers, ontologists and metaphysicians, make conscious ontological claims, and they might make such claims even if ordinary speakers would not make them or would make them only tacitly. Also scholars of other disciplines might make ontological claims even if ordinary speakers did not. In any case it is irrelevant to ask as Azzouni does whether ordinary uses of "there is" in English are understood by speakers to carry ontological commitment. We should instead ask whether uses of the Latin word "Est" were understood by Wolff to carry ontological commitment.

The question of ontological commitment can be raised with respect to statements belonging to or theories formulated in three different languages. We can ask for the ontological commitments of statements belonging to the language of first-order predicate logic, of statements belonging to the language of some other logic such as higher order logic or modal logic, and of statements belonging to a natural language such as English. It is very important to distinguish these cases from each other.

The criterion of ontological commitment formulated by Quine is concerned directly only with the ontological commitments of statements belonging to predicate logic. Indeed, Quine mostly only considers the commit-
ments of statements belonging to first-order predicate logic. However, Quine is clear that the criterion is also relevant to the ontological commitments of at least some statements made in other languages.

In these cases, however, it must be used indirectly; we must first translate the statements of other languages to statements of first-order predicate logic and then find the ontological commitments of the translation. As Quine says if we prefer another language form, for example that of combi-
tors, we can still bring our criterion of ontological form to bear in so far as we are content to accept appropriate systematic corre-
lations between idioms of the aberrant language and the familiar language of quantification.

93 It seems to me that what Quine says taking the language of combinators as one example can also be applied to natural languages, even if Quine himself had not intended such an application, especially if we consider Montague’s claim that natural languages can be treated as formal languages. Indeed, natural languages are far closer to the language of combinators than to the language of standard predicate logic, as natural languages have no variables in the same sense as predicate logic.94 In the case of natural languages such a translation of their statements implies formalizing them. However, it is

93 Quine considers it possible that it might happen relative to a really weird language that we cannot make any sense of ontological commitment. He even goes so far as to say that entity and objectuality might not be applicable to every conceptual scheme, being provincial traits of our conceptual scheme. Here he seems to go overboard into an unsustainable relativism.

94 It has indeed often been pointed out that the pronouns and other proforms of natural languages have some similarities to variables. Such similarities are real, and important; I think it follows from them that the use of proforms such as pronouns in natural languages usually expresses ontological commitment. Therefore such similarities enable us to expand our criterion of ontological commitment. They enable us to say that if a statement contains an expression which can be be referred back to with an anaphoric proform then the statement must also be committed to entities denoted by that expression or falling into its extension even if the statement does not contain any explicit quantifier. This is vitally important in finding out the ontological commitments of expressions in natural languages. Nevertheless, such similarities must not be overemphasized. There are also equally crucial differences. Most natural languages can make a quantified statement without using any proform, while predicate logic cannot make any quantified statement without using variables. E. g. we can just say ”There is a man.”, in which sentence there are no proforms, without having to say anything like ”There is something such that it is a man.”, which would be analogous to the way the sentence is expressed in the language of predicate logic. It follows from this lack of variables that even when a natural language sentence is committed to something, it is not always clear exactly what it is committed to. This suggests we should be very cautious in assessing the commitments of a sentence before we can symbolize it exactly.
not only this that it implies. A mere formalization is not a translation, as
two statements with the same form might yet have a different meaning and
if neither of them is logically true they can even have different truth-values.
A statement of natural language can only be translated to a language of
predicate logic whose non-logical constants have a fixed interpretation. Of
course, getting a formalization is big step towards a translation, since once
you have a formalization you only need a determinate correspondence be-
tween the constants of the two languages. How such an interpretation is
given or can be given is not relevant here. All that matters is that it can
be given. In his later relativistic phase Quine actually doubts if such an
interpretation can be at all be given. However, from a pragmatic point of
view which Quine often professes to follow such relativism is utterly inde-
fensible. Ultimately knowledge of logical truths would be useless unless it
could be united to a knowledge of non-logical truths. After all, as knowl-
edge of a logically true statement cannot in itself help us to bake bread or
heal disease or in short be of any use, knowledge of what is logically true is
useful only because it enables us to infer one non-logically true statement
from another, i.e. pass from knowledge of one non-logically true state-
ment to the knowledge of another non-logically true statement. However, if
statements have no fixed meanings or even truth-values there surely cannot
be any knowledge of them. If the translation is adequate, then the com-
mitments of the translations are also the ontological commitments of the
original statements. More specifically, so far as the translation is adequate,
so far can we say adequately what the commitments of the original state-
ment are. Whatever imperfection or uncertainty attaches to the translation,
so much imperfection uncertainty attaches also to the judgement concerning
ontological commitments attained by the means of the criterion.

If a natural language assertion is ambiguous or vague then we can say
that it is committed to some entities simpliciter only in the supervaluationist
sense, i.e. if all precise symbolizations of it are committed to them. As
the supervaluationists speak of supertruth in the case of a statement that
is true relative to all ways of making it precise, i.e. all precisifications of
it, we can analogously speak of ontological supercommitment. An assertion
carries ontologically supercommitment to some entities iff all precisifications
of it carry ontological commitment to them\(^\text{95}\). If the formalizations of an

\(^{95}\text{Other theories of vagueness can also be applied to the question of the ontological}

assertion disagree in what they are committed to we can only say that the assertion is committed to some entities relative to a reading of it, or that the speaker is ontologically committed in a pragmatic sense to some entities because of the way he intends his statement to be understood.

What Quine himself says about the ontological commitments of ordinary languages is not wholly satisfactory. Quine says that

In a loose way we often can speak of ontological presuppositions at the level of ordinary language, but this makes sense just so far as we have in mind some likeliest, most obvious way of schematizing the discourse in question along quantificational lines.

Quine gives no here reason why we should have in mind any obvious way of formalizing the discourse, but this claim obviously follows from his famous thesis of the indeterminacy of translation which is connected with his deflationary theory of truth. Quine’s claim is of course analogous to the claim that in order to speak meaningfully about the truth-conditions of the sentences of a natural language we would have to have in mind some obvious way of schematizing those sentences in predicate logic. This is a claim that follows Quine’s deflationary theory of truth; however, a follower of an (even weakly) inflationary correspondence theory of truth can deny it and similarly a follower of a more inflationary theory of ontological commitment can deny the the corresponding claim concerning ontological commitments.

The reference to a systematic correlation in the earlier quote seems far more perspicuous to me. However, that a correlation is systematic surely does not imply that it would be obvious or at all easy to find out. All that matters that there is some systematic way of formalizing the discourse, no matter how difficult to find out or how cumbersome to apply and no matter whether anyone has yet found one. What does systematicity mean? It means that there is a general rule enabling us to correlate any number of expressions of one language with those of the other. I suggest that this is the same thing as recursivity.

Of course there can be difficulties about what is the best way to translate or even formalize an assertion or theory expressed in a natural language.

commitments of a natural languages. If e. g. it makes sense to speak of degrees of truth, which is highly controversial, it might also make sense to speak of degrees of ontological commitment.
Quine’s and Church’s criteria are of course not meant to solve all problems connected with ontological commitment at once; surely no criterion that can be expressed succinctly can do that. Despite this the criteria can be useful in finding out the ontological commitments of a statement. Finding out how to translate a natural language sentence into the language of predicate logic (enriched with non-logical constants) can surely not be a wholly impossible task, as any student taking a basic course in logic is supposed to be capable of doing it in many cases! However, all of Quine’s opinions about what formalization involves are not to be taken as part of his criterion of ontological commitment. One can accept Quine’s criterion while differing about these questions. E. g. Quine thought that formalization could involve some revision of the formalized theory, but one can consistently accept Quine’s criterion while holding that formalization should not involve any alteration of what is formalized.

Some principles of formalization should be obvious, and are very widely accepted. A formalization of the sentences of a natural language should

1. be recursive

2. capture as many intuitive entailments as possible and

3. be as simple as it can be consistently with the two earlier requirements.

First of all, the formalization of natural language sentences should ideally be recursive, so that we can be sure that every sentence can in principle be formalized. The recursivity of natural languages follows, as is well known, from the requirement that natural languages should be learnable. This recursivity implies that we must strive towards a fully mechanical rule of formalization, that does not require the possession of any linguistic intuitions in its applier and hence could be applied by computers as well as by men. However, unlike what is often suggested, recursivity need not imply full compositionality. Josh Dever has in [Dev08] distinguished two constraints implied by compositionality; semantic closure and semantic locality. According to the first only semantic information can go into determination of the semantic value of a complex expression. According to the second only information derived from parts of a complex expression can go into the determination of the semantic value of the expression. The latter constraint implies that the semantic rules must be context-free. Recursivity does not
by itself imply either of these two constraints. It allows that in formalizing a complex expression we can take into account directly its subexpressions and not merely their semantic values, so it does not imply semantic closure. It also allows us to take the semantic values of previous and succeeding expressions into account in determining the value of an expression, so it does not imply semantic locality either.

However, a semantically local formalization is as such, other things being equal, simpler than one that is not semantically local, so it may follow from recursivity together with the third principle that we should strive towards a semantically local formalization unless of course other things are not equal (as they often are not).

It follows from recursivity together with simplicity that expressions with the same syntactic form (or forms, if the expression is syntactically ambiguous) should usually be formalized so that they end up having the same logical forms when they occur in the same context. Perhaps the only exception is that we can give a phrasal expression like "kick the bucket" an additional formalization on which it is formalized as an atomic predicate (in this example the same as the one symbolizing the verb "kill") in addition to the formalization following from its composition. However, it follows from the requirement that natural language be learnable that there are only a finite number of such phrasal expressions. It is not consistent with learnability and recursivity that there should be a potentially infinite number of complex expressions all of which would be given formalizations separately. Of course we can give many formalizations, representing many readings, to a structurally ambiguous sentence. Usually many of the semantically possible readings of a sentence are of course pragmatically excluded; however, formalization need not take such pragmatic matters into account but should only deal with all semantically possible meanings that are determined conventionally.

Secondly the formalization should make valid inferences that are intuitively valid, i.e. inferences that most people and the formalizer himself would be willing to make in the course of everyday life. This leads room for revision of ordinary practice, as the formalizer can decide not to respect a commonly accepted practice if he himself decides not to apply it, but not very much room.

Thirdly, we should use the simplest possible formalization that respects
the other constraints. Simplest here means the simplest overall; the simplest formalization of a single sentence might because of recursivity require an excessively complex formalization of other sentences, and therefore have to be rejected. Again it is not always clear what formalization is the simplest; however, there are cases where it is intuitively unmistakable that one formalization is simpler than another.

Chihara asks in [Chi68, page 38] whether \((\exists x)(x \text{ is a winged horse})\) is ontologically committed to wings. This question is not very clear as it stands; this curious mixture of words and symbols is not a well-formed expression of either English or of predicate logic, so it has no ontological commitments. It could be interpreted in various ways so that it would carry ontological commitment, but it might receive either an interpretation according to which it was committed to wings and one according to which it was not. Presumably Chihara really wants to ask whether the sentence "There is a winged horse" is committed to wings. The answer depends on how the sentence should be formalized. Should it be formalized as \((\exists x)(W(x))\) or as \((\exists x)(H(x) \& (\exists y)(W(y) \& H(x,y)))\)? The very weak theory of formalization I have suggested above suffices to give a quite determinate answer to this question.

As the phrase "winged horse" should because of recursivity and simplicity be formalized in the same way in all sentences, the answer to this depends on whether the inference from the statement "Perseus rode a winged horse" to the statement "Pegasus rode a horse and it had wings." is intuitively valid. However, this question should be easy to answer on the basis of the second principle, that formalization should make intuitively valid inferences logically valid. The crucial question is the following: if you were making inferences about Greek mythology or if you believed mythology to be history would you be willing to make the inference? If your history teacher tells you repeatedly in class: "Perseus rode a winged horse." and asks you in an examination "Did a horse Perseus rode have wings?", and you are eager to pass the examination, would you answer that he did? Most people would surely answer affirmatively to this question! Therefore the inference is intuitively valid for them. However, the validity of this inference can surely be most easily explained by assuming that the first sentence has the symbolization \((\exists x)(H(x) \& (\exists y)(W(y) \& H(x,y)))\), where "W" symbolizes the predicate of being a wing, the dyadic predicate "H" the predicate of having
and the monadic predicate "H" the predicate of being a horse. Therefore all people that answer affirmatively to this question should accept that this is the correct formalization. I will suppose that the reader is among them. This sentence, however, implies logically by standard first order logic the sentence \((\exists x)(W(x))\), which is committed to wings, and therefore according to the criterion of ontological commitment I have proposed the original sentence "There is a winged horse." also carries (implicit formal) ontological commitment to wings.

Therefore the criterion should be applied to sentences after they have been semantically analyzed. However, Chihara opposes the view that the criterion should be applied to "fully analyzed" sentences. I agree of course that the criterion cannot be applicable only when we have reached a complete semantical analysis, since as Chihara suggests reaching such an analysis may not be feasible. However, my conclusion does not imply that we would not be now in position to apply the criterion but should have to await the construction of the required semantical theory. My suggestion implies that we should analyse the sentence as far as we can at the present time and then apply the criterion. It is then of course always possible that when we achieve better semantical analyses we will have to revise our assessments of ontological commitment. We certainly have well-developed semantical theories making use of decompositional analysis right now (e.g. Dowty’s version of Montague semantics in [Dow79], Parsons’s version of event semantics in [Par94], etc.) which go far below surface syntax in analysis even if none of them is anywhere near complete. If we are to consider the ontological commitments of natural languages at all usefully it must be as part of such grand though necessarily incomplete projects in lexical semantics. It may be that none of these theories is very certain either; however, this is as it has to be. It cannot be demanded of an adequate criterion of ontological commitment - or an adequate criterion of anything - that it could lead us to wholly certain results, results that need never be revised. We would have no criteria for anything if we made such a demand! Fallible criteria can be quite useful.

Chihara also objects that interpreted in this way the criterion is not in the spirit of Quine as a semantic atomism he would oppose is implicit in it. I have already pointed out, however, that all of Quine’s opinions concerning semantics should not be taken to be a part of his criterion. If
this interpretation of the criterion is inconsistent with Quine’s semantical holism, then so much the worse for semantical holism!

Chihara prefers an interpretation suggested by Osvaldo Chateaubriand according to which the notion of ontological commitment should be defined relative to a metatheory. Chateaubriand’s suggestion surely fits in well with Quine’s later theory of ontological relativity, as Chihara says in [Chi74, page 80], though it is less clear if it expresses Quine’s original intention in formulating his criterion. Even Quine’s earlier theory had relativistic tendencies, however, so it might well do so.

In any case, Chateaubriand’s suggestion is not a good interpretation if we want an interpretation of the criterion that not only respects Quine’s intentions but is independently defensible as an adequate criterion of ontological commitment. It clearly expresses a radically deflationary theory of ontological commitment similar to radically deflationary theories of truth (like those of Quine himself, the later Field and Brandom), and can be argued against in a way perfectly analogous to the way deflationary theories of truth are argued against. Relativisation of semantical notions to a metatheory leads to their complete trivialization and to absurd anti-realism, as I have already argued in the case of the notions of truth and denotation. Chateaubriand’s suggestion would similarly lead to the extremely absurd conclusion that a theory can be correctly taken to be committed to anything if we are working in a suitable metatheory, and no such statement of ontological commitment would be more correct than another.

According to Chihara’s exposition of Chateaubriand’s theory in [Chi74, page 79] S carries ontological commitment to an entity x such that F relative to a metatheory T if and only if for some open sentence M (\(\exists x)(M)\) is an assertion of S and (\(x)(M \rightarrow F)\) is an assertion of T. Chihara suggests further that community-wide beliefs provide us with a likely candidate for the overall home theory relative to which we ask about the ontological commitments of theories we specify.

This clearly leads to absurd consequences. Let us suppose that the community in question is the community of evolutionary biologists and their overall home theory is the theory T of evolutionary biology and that they are trying to find out the ontological commitments of the creationist theory S of Christian fundamentalists. Now if M means that x is a man and F that x is a man descended from apes, then (\(x)(M \rightarrow F)\) is an assertion of T. If S
is the creationist theory then it surely includes the assertion that there are
men. Thus \((\exists x)(M)\) is an assertion of \(S\). It follows from Chateaubriand’s
interpretation of Quine’s criterion that \(S\) carries commitment to men de-
sceded from apes relative to \(T\). Therefore evolutionary biologists would
have to conclude that creationists are committed to men descended from
apes. We can show similarly that fundamentalists should conclude that
evolutionary biologists are committed to men descended from a specially
created Adam and Eve. These conclusions are of course obviously absurd,
as creationists are only committed to men descended from Adam and Eve
and evolutionary biologists should surely see this.

Perhaps Chihara would not admit that the community of evolutionary
biologists or the community of fundamentalists are big enough that their
overall theories could function as an overall home theory. Chihara does
speak following Quine of ”the whole linguistic community” so perhaps he
supposes that his theory only applies to so big communities as all speak-
ers of English or all speakers of German or even speakers of any language
translatable to English (though evolutionary biologists surely share a com-
mon technical jargon, so they might also be called a linguistic community).
However, it is not clear that there factually are any beliefs shared by so big
communities. For any belief some lunatic might be found who would quite
sincerely deny it, even including beliefs commonly taken to be analytic.
Thus this interpretation of Chihara’s notion of communities would trivial-
ize Chateaubriand’s criterion completely. Besides, we could assume for the
sake of argument that in the future all speakers of English would believe
in evolutionary theory and that they would ask for the commitments of a
past creationist theory no one held any more (or even a purely hypothetical
creationist theory). In this case Chihara’s interpretation would definitely
predict that such future speakers of English should assume that creation-
ists were committed to men descended from apes. However, of course they
would be wrong in making such an assumption even in that case.

Chihara himself expresses in [Chi68, page 42] doubts whether such a
relativised notion of ontological commitment would have any serious use
in philosophy. Surely we ought not be satisfied with such a criterion; the
important question is not what Quine himself thought or would have thought
but what criterion is correct and useful. The interpretation of the criterion
I have suggested, even if it leads to open problems and uncertainty, at least
does not lead to absurdity or complete triviality.

Nevertheless Chihara’s and Chateaubriand’s theory is not wholly misguided, and with small modifications it can be made to yield the same conclusions as the one I have proposed. It is not hard to see how Chateaubriand’s theory could be corrected. Its error is that any assertion of the form $(x)(M \to F)$ in the metatheory can be used in seeking for ontological commitment. Some selection between different kinds of assertions of the metatheory additional to their syntactic form is necessary. Obviously in cases where $(x)(M \to F)$ is a theorem of predicate logic, e.g. where $M(x)$ is the same as $F(x) & G(x)$, the assertion can be used. Also if $(x)(M \to F)$ is a stipulative definition made explicitly by the formulator of the metatheory it can be used. However, one would also like to use formulas of the metatheory that are meaning postulates for the object language. Quine would of course not be happy with this, but others who have advocated a similar criterion such as Church would have no reason to object to it. It is in any case necessary if the criterion is to be really useful.

5.16 Materially Implicit Ontological Commitments

However, restricting the implication in question to logical implication in the narrow sense would still in my view make the criterion too strict. Some kind of analytical entailment should in my view also be allowed. Quine himself of course would not have allowed it, as he distrusted the notion of analyticity. However, since Quine’s arguments against analyticity are considered unsound by many philosophers, this need not be a sufficient reason to restrict the implication to logical implication. Many of us would surely be inclined to say, at least on some occasions, that if a theory is committed to bachelors it is committed to unmarried human beings and that if a theory is committed to red objects then it is committed to coloured objects. Also as Chihara pointed out, Quine himself would have wanted to say that ‘$(\exists x)(x \text{ is a set})’ carries ontological commitment to an entity $x$ such that $x$ is an abstract entity, and this requires some kind of analytical containment between the concept of a set and the concept of an abstract entity. Let us call such wider concepts of ontological commitment implicit material ontological commitment. The use of this wider concept of ontological commitment need not preclude us also from using the narrower concept; one concept is useful
in some contexts and the other is useful in other contexts. We can make use of both so long as we do not confuse them with each other.

The first example I mentioned can be accommodated rather simply if we have a good understanding of what stipulative definitions are. Definitions are often in modern logic taken to be metalinguistic abbreviations. However, this understanding of definition is in no way necessary, and does not correspond to the way the word is used outside logic. There have been logical systems (such as Lesniewski’s systems) in which definitions have been taken to genuinely enrich the object language syntactically by introducing new expressions to it, though of course a definition should not enrich a language semantically, should not increase the expressive power of a language. Thus if \( L \) is a language with an interpretation \( I \), then we can expand it to a language \( L' \) with an interpretation \( I' \) by means of definitions. For example in the case of predicates definitions are of the kind \( P_i(t_1,\ldots,t_n) \equiv_{\text{def}} p_1 \land \cdots \land p_n \), i. e. functions from the set of formulas of the language \( L \) to predicates of the language \( L' \) where the formulae \( p_1,\ldots,p_n \) are formulae of \( L \) and where predicate \( P_i \) (the definiendum) belongs to \( L' \) but not to \( L \). We then expand \( I \) to a new interpretation \( I' \) so that every sequence of individuals belongs to the extension of \( P_i \) iff it satisfies all the formulae \( p_1,\ldots,p_n \), i. e. for all terms \( t_1,\ldots,t_n \) of \( L \) and all assignments \( g, \models_{I'} P_i(t_1,\ldots,t_n) \equiv \models_{I,g} p_1 \land \cdots \land p_n \).

Using this understanding of definitions we can then give the following rough initial definitions of narrow analytical ontological commitment (where narrow analyticity is the kind of analyticity defined by Frege).

**Metadefinition 9** A sentence \( S \) carries narrow explicit analytical ontological commitment to entities \( \alpha \) such that \( M \) if and only if there is a sentence \( S' \) s. t. \( S' \) carries explicit ontological commitment to entities \( \alpha \) such that \( M \) and \( S' \) can be derived from \( S \) be repeatedly replacing definienda with their definienses or conversely.

**Metadefinition 10** A sentence \( S \) carries narrow implicit analytical ontological commitment to entities \( \alpha \) such that \( M \) if and only if there is such a sentence \( S' \) that \( S' \) carries formal implicit ontological commitment to entities \( \alpha \) such that \( M \) and \( S' \) can be derived from \( S \) by repeatedly replacing definienda with their definienses or conversely.

However, this narrow analytical commitment, whether explicit or implicit, may not be the only kind of implicit ontological commitment. Many
of us would naturally say that if a theory is committed to red objects then it is committed to coloured objects. However, this need not hold of narrow analytical ontological commitment, since arguably (at least phenomenal) redness cannot be defined in terms of colour. Peacock draws attention in [Pea11, page 88] to an even more gripping example: the implicit commitments of a theory containing the sentences “there are some people who are taller than some other people” should include people who are shorter than some other people, for such people would have to exist, given that the theory is true.

The problem of how to explicate such a notion of implicit material ontological commitment is very hard, and I do not have any final solution to offer. Perhaps the best way of dealing with such materially implicit ontological commitments is to make use of Carnap’s theory of meaning postulates in [Car52], as amended by Kemeny in [Kem95], who replaced Carnap’s state descriptions with models. Formally we introduce a new set to the metatheory, the set of meaning postulates, which I will symbolize with $\mathbb{P}$. This can be taken to be just a set of sentences; e. g. if $T(x, y)$ symbolizes the statement that $x$ is taller than $y$, and $S(x, y)$ the statement that $x$ is shorter than $y$, then the sentence $(\forall x)(T(x, y) \equiv S(y, x)) \in \mathbb{P}$ We can just take the Metadefinition 8 and restrict the quantification in it to quantification over those models in which the meaning postulates hold, where meaning postulates $\mathbb{P}$ hold in a model $\mathcal{M}$ iff for all $\phi \in \mathcal{M}$ and all assignments $g \models_{\mathcal{M}, g} \phi$. Quine himself would have disliked this idea very much because of his doubts concerning analyticity, but this should be no argument against extending the theory of ontological commitments beyond Quine’s own thoughts in this way, especially if Quine’s argument against analyticity was based, as I argued, in an unacceptable verificationism and led to an unacceptable semantic holism.

A rough formulation of such a theory will then be the following:

**Metadefinition 11** An assertion $\phi$ carries broad analytic implicit ontological commitment to entities $\alpha$ such that $M$ relative to assignment $g$ if and only if there is such a sentence $\psi$ that $\psi$ carries explicit ontological commitment to entities $\alpha$ such that $M$ relative to $g$ and for all models $\mathcal{M}$ in which the meaning postulates $\mathbb{P}$ hold and for all assignments $h \phi$ is true with respect to $\mathcal{M}$ and $h$ only if $\psi$ is true with respect to $\mathcal{M}$ and $h$ (i. e. $\models_{\mathcal{M}, h} \phi \rightarrow \models_{\mathcal{M}, h} (\exists x)\psi$).
5.17 A Model-theoretical Account of Ontological Commitment

In modern model theory the kind of absolute notion of truth which was used in Tarski’s model theory is replaced with a notion of truth relativised to interpretations and domains. I have already been making use of such a relativisation of truth as an auxiliary concept in defining formally implicit ontological commitment; however, I have not yet relativised any of the notions of ontological commitment which are my main subject. However, it is of course possible to relativise the notion of ontological commitments similarly to the notion of truth. We can speak of the ontological commitments of a statement relative to an interpretation of its non-logical constants. We can thus say that relative to a model whose interpretation function associates the word ”man” with asses, the sentence ”There are men.” would be ontologically committed to asses.

In current model theory truth is in fact relativised to two factors; to the interpretation of non-logical constants and to an individual domain. These relativisations are to a large extent independent of each other and we can consider them separately. In fact Tarski himself originally understood models in [Tar83b, page 417] in a radically different way from the way the concept of a model is usually employed in modern metamathematics. John Etchemendy has noticed this difference between Tarski’s concept of a model and the standard one in [Etc88]. Tarski says that a model of a class of sentences is a sequence of objects that satisfies every sentential function derived from those sentences by replacing their extra-logical concepts with variables. Thus a model in Tarski’s view only gives an interpretation to non-logical constants; it does not provide any domain for quantification.

They are not usually treated as wholly independent, as it is usually assumed that the interpretation of a singular term must be a member of the domain of the model and the interpretation of a predicate a subset of the domain. However, this assumption is sometimes dropped in non-standard semantics, for example in some semantics for some systems of positive free logic, where the interpretation of a term is allowed to be outside the domain of quantification. It seems to me that we can well drop this assumption, even if our logic is not a positive free logic; we can avoid accepting the axioms of positive free logic by altering the truth-conditions slightly even if some interpretations are allowed to be outside the realm of quantification (e. g. we could just degree atomic sentences the denotations of whose terms are outside the domain to be false and so gain the axioms of negative free logic.). If this is done, the two relativisations become wholly independent, and ontological commitment becomes wholly independent of the domain of the model.
It might seem that this is not a novel idea, for John Bacon has already tried to relativise ontological commitment to models in [Bac87]. However, Bacon does not really define the truth of a theory (or sentence) relative to a model as I will do; rather he defines the ontological commitments of a model without referring to any theory, and then tries to define the ontological commitments of a theory by using the notion of a model as an auxiliary concept, without relativising the ultimate definiendum into models. Bacon says in [Bac87, page 1] that his basic explicandum is existence in a model; Bacon symbolizes existence in a model \( m \) as \( O(m) \). Bacon then says that the explicatum he proposes is first approximated as membership in the model’s domain. This is indeed a natural explicatum of that explicandum; however, this explication is rather trivial and has little to do with the interesting concept of the ontological commitments of theories or sentences. What Bacon does is rather like trying to define truth in a model but not the truth of any sentence but truth simpliciter, which would obviously be senseless.

Bacon’s explication of the ontological commitments of a theory fails in an instructive way. According to Bacon’s initial definition \( F \)’s exist according to \( T \) iff \( F \)’s exist in every intended model of \( T \). According to Bacon this means that \( F \)’s exist according to \( T \) iff \( F \cap O(m) \neq \Lambda \) for all \( m \in M_t \), and \( O(T) = \{ F : F \)’s exist according to \( T \} \). However, Bacon notes that this definition does not allow a theory to commit us to something that in fact doesn’t exist, e. g. it does not allow a theory affirming \(( \exists x)(x \text{ is a unicorn})\) to commit us to unicorns. However, this means that the theory fails radically.

Bacon tries to correct this defect by relativising ontological commitment to subjects. This is intuitively a wrong way to solve the problem; if a theory says that there are unicorns, then this suffices to say that the theory carries commitment to unicorns; we do not have to know who has proposed the theory. Bacon is here confusing the semantic concept of ontological commitment with the pragmatical concept of ontological commitment. However, his definition of personal commitment does not succeed very well even in defining the pragmatic concept of ontological commitment. Bacon says in [Bac87, page 6]:

\[
T \text{ commits } S \text{ to } F \text{’s iff } S \text{ is committed to claim of } T \text{ that } : T \text{ is true and } F \in O(T).
\]

This definition has a strong whiff of circularity; the notion of commit-
ment occurs both in the definiendum and in the definiens! One might try to defend Bacon’s definition by saying that the two concepts of commitment are not quite the same; the one in the definiens is a commitment to (the truth of) a claim and the one in the definiendum commitment to the existence of entities. Nevertheless, the two concepts are so similar to each other that the definition does not seem to be very illuminating. Besides, it is not very clear how the notion of commitment that occurs in the definiens is to be understood. If it were understood so that commitment to claiming is just the same as being disposed to claim then the definition would have unacceptable consequences; it would imply that a theory could not commit a person to the existence of any entities unless the person knew enough semantics and model theory to be capable of making the rather complex claim that $F \in O(T)$. Perhaps the notion can be understood in some other way so that it does not have such consequences, but Bacon does not give us any reason to think that it can be understood so that the definition is neither circular nor has these unacceptable consequences.

The failure of Bacon’s theory can be explained by pointing out that Bacon’s interpretation of Quine’s theory is defective. However, it would be uncharitable to say blankly that Bacon has misunderstood Quine’s theory. As I have already pointed out, many of Quine’s formulations are obscure and inconsistent, so Bacon may indeed have captured one possible interpretation of Quine’s theory; however, it is an interpretation that would make the theory a very bad theory. A better interpretation can be found. Bacon says in [Bac87, page 2] that to be is to be the value of a bindable variable. However, though some of Quine’s formulations may suggest this, it already trivializes the theory of ontological commitment, since it makes all theories with the same language to have the same ontological commitments. Bacon then commits the same mistake we have seen Bergmann and Carnap committed. It is more accurate to say that to be is to be the value of a variable that is actually bound, and to be even more accurate we must add to be actually bound by an objectual quantifier in a true sentence. Generalizing this to false theories, to be according to a theory $T$ is to be the value of a variable actually bound by an objectual quantifier in a sentence that the theory $T$ contains or implies if the theory is true. However, even these formulations are not unambiguous; no formulation in natural language is, and the use of logical symbolism is necessary to get more accurate formulations.
Bacons’ theory can then provide no guidance in our attempt to define the notion of ontological commitment relative to models, but we must start from scratch.

Before actually defining ontological commitment relative to interpretations or models, we must redefine objectuality and non-objectuality for quantifiers in languages in which only the logical constants have a fixed interpretation. I will use \( \| t_i \|_{I,g} \) to stand for the semantic value of (singular or predicate) term \( t_i \) relative to interpretation \( I \) and assignment \( g \). If \( t_i \) is a constant, then \( \| t_i \|_{I,g} = I(t_i) \) while if \( t_i \) is a variable, then \( \| t_i \|_{I,g} = g(t_i) \).

**Metadefinition 12** An unrestricted existential quantifier \( (\exists x) \) is objectual with respect to the interpretation \( I \) iff for any formula \( \phi \) and for any assignment \( g \), \((\exists x)\phi\) is true with respect to \( I \) and \( g \) if and only if there is such an \( \alpha \) that \( \phi \) is true with respect to \( I \) and \( g(\alpha/x) \).

**Metadefinition 13** An operator \( (\exists i) x \) is an (unrestricted) non-objectual existential quantifier with respect to the interpretation \( I \) if and only if for all formulas \( \phi \) there is such an interpretation \( J \) that for all assignments \( g \), \((\exists i) x \phi \) is true with respect to \( I \) and \( g \) if and only if there is such an entity \( \alpha \) that \( \phi \) is true with respect to the interpretation \( J \) and the assignment \( g(\alpha/x) \) and for some formula \( \phi \) \( \phi \) is true with respect to \( J \) and the assignment \( g(\alpha/x) \) but is not true with respect to \( g \) and every atomic formula \( p \) is true with respect to \( J \) and the assignment \( g(\alpha/x) \) iff it is true with respect to \( I \) and the assignment \( g(\alpha/x) \).

So if \((\exists x)\) is objectual, it is not necessary or even enough for the truth of \((\exists x)\phi \) with respect to \( I \) and \( g \) that for some term \( a \), \( \phi(a/x) \) would be true with respect to \( g \) and \( g \).

We must again begin from the case where an atomic sentence is in the scope of an existential quantifier.

**Metadefinition 14** Sentence \( (\exists x)R(t_1, \ldots, x, \ldots, t_n) \) carries explicit ontological commitment to entities \( \alpha \) such that \( Q(\beta_1, \ldots, \alpha, \ldots, \beta_n) \) relative to interpretation \( I \) and assignment \( g \) iff \( \| R \|_{I,g} = Q, \| t_1 \|_{I,g} = \beta_1, \ldots, \| x \| = \alpha, \ldots \) and \( \| t_n \|_{I,g} = \beta_n \) and the quantifier \( (\exists x) \) is objectual.

The next case is that where negative sentences are in the scope of the existential quantifier.
Metadefinition 15  Sentence $\phi$ carries explicit ontological commitment to entities $\alpha$ such that it is not the case that $M$ relative to interpretation $I$ and assignment $g$ if and only if it is identical with a sentence $(\exists x)p$ and sentence $(\exists x)\neg p$ carries ontological commitment to entities $\alpha$ such that $M$ relative to interpretation $I$ and assignment $g$.

The case of conjunction is simple.

Metadefinition 16  Sentence $\phi$ carries explicit ontological commitment to entities $\alpha$ such that $M$ and $N$ relative to interpretation $I$ and assignment $g$ if and only if it is identical with a sentence $(\exists x)(p \& q)$ and the sentence $(\exists x)p$ carries ontological commitment to entities $\alpha$ such that $M$ relative to $I$ and $g$ and the sentence $(\exists x)q$ carries ontological commitment to entities $\alpha$ such that $N$ relative to $I$ and $g$.

Next follow sentences where a disjunction is in the scope of the existential quantifier.

Metadefinition 17  Sentence $\phi$ carries explicit ontological commitment to entities $\alpha$ such that $M$ and $N$ relative to interpretation $I$ and assignment $g$ if and only if it is identical with a sentence $(\exists x)(p \lor q)$ and the sentence $(\exists x)p$ carries explicit ontological commitment to entities $\alpha$ such that $M$ relative to $I$ and $g$ and the sentence $(\exists x)q$ carries ontological commitment to entities $\alpha$ such that $N$ relative to $I$ and $g$.

Finally we must consider quantified sentences.

Metadefinition 18  Sentence $\phi$ carries explicit ontological commitment to entities $\alpha$ such that for some entity $\beta$ $M$ relative to interpretation $I$ and assignment $g$ if and only if it identical with a sentence $(\exists y)(\exists x)p$ and for all entities $\beta$ the sentence $(\exists x)p$ carries explicit ontological commitment to entities $\alpha$ such that $M$ relative to interpretation $I$ and assignment $g(\beta/y)$.

Next we must consider what happens when individual domains are added to models. One might think that the relativisation to an individual domain does not affect the notion of ontological commitment in any interesting way, even though it affects the notion of material truth considerably, since the entities a statement is ontologically committed to need not exist and hence
a fortiori need not be members of the domain in any model. However, this can also be thought of in another way, which makes the second sort of relativisation relevant. Just because of the hyperintensional character of the notion of explicit ontological commitment, even if some entities do not belong to the domain of the model, we can be explicitly committed to entities that are among them yet do belong to the domain of the model.

The concepts of ontological commitment relativised to models can be defined exactly as follows below. Here \( \|t_i\|_{\mathfrak{M},g} \) is the semantic value of (singular or predicate) term \( t_i \) relative to model \( \mathfrak{M} \) and assignment \( g \). If \( \mathfrak{M} = \langle I, D \rangle \) and \( t_i \) is a constant, then \( \|t_i\|_{\mathfrak{M},g} = I(t_i) \) while if \( t_i \) is a variable, then \( \|t_i\|_{\mathfrak{M},g} = g(t_i) \).

The definition of the objectuality of quantifiers in languages in which only logical constants have a fixed interpretation and all quantifiers are restricted to some domain is now an obvious modification of its definition for languages all of whose constants have a fixed interpretation.

**Metadefinition 19** An existential quantifier \((\exists x)\) is objectual with respect to the model \( \mathfrak{M} = \langle I, D \rangle \) iff for any formula \( \phi \) and for any assignment \( g \), \((\exists x)\phi\) is true with respect to \( \mathfrak{M} \) and \( g \) if and only if there is such an \( \alpha \in D \) that \( \phi \) is true with respect to \( \mathfrak{M} \) and \( g(\alpha/x) \).

**Metadefinition 20** An operator \((\exists_i x)\) is a non-objectual existential quantifier with respect to the model \( \mathfrak{M} = \langle I, D \rangle \) if and only if for all formulas \( \phi \) there is such a model \( \mathfrak{N} \) that for all assignments \( g \), \((\exists_i x)\phi\) is true with respect to \( \mathfrak{N} \) and \( g \) if and only if there is such an entity \( \alpha \in D \) that \( \phi \) is true with respect to the model \( \mathfrak{N} \) and the assignment \( g(\alpha/x) \) and for some formula \( \phi \) \( \phi \) is true with respect to \( \mathfrak{N} \) and the assignment \( g(\alpha/x) \) but is not true with respect to \( \mathfrak{M} \) and \( g \) and every atomic formula \( p \) is true with respect to \( \mathfrak{M} \) and the assignment \( g(\alpha/x) \) iff it is true with respect to \( \mathfrak{N} \) and the assignment \( g(\alpha/x) \).

Any quantifier which is given a model-theoretic interpretation is naturally a restricted quantifier, a quantifier restricted to the domain of the model. Unrestricted quantifiers cannot be treated in model theory. Therefore the notion of ontological commitment relativised to models is a significantly modified notion of ontological commitment.

Thus it is not necessary or even enough for the truth of \(((\exists x)\phi\) with
respect to $\mathcal{M}$ and $g$ that for some term $a$, $\phi(a/x)$ would be true with respect to $\mathcal{M}$ and $g$.

We must again begin from the case where an atomic sentence is in the scope of an existential quantifier.

**Definition 7** A sentence $\phi$ carries explicit ontological commitment to entities $\alpha$ such that $Q(\beta_1, \ldots, \alpha, \ldots, \beta_n)$ and $\alpha \in D$ relative to model $\mathcal{M} = \langle I, D \rangle$ and assignment $g$ if and only if it is identical with such a sentence $(\exists x)R(t_1, \ldots, x, \ldots, t_n)$ that the existential quantifier $(\exists x)$ is objectual with respect to the model $\mathcal{M} = \langle I, D \rangle$ and $\|R\|_{\mathcal{M}, g} = Q$, $\|t_1\|_{\mathcal{M}, g} = \beta_1$, $\ldots$ and $\|t_n\|_{\mathcal{M}, g} = \beta_n$.

The next case is that where negative sentences are in the scope of the existential quantifier.

**Metadefinition 21** Sentence $\phi$ carries explicit ontological commitment to entities $\alpha$ such that it is not the case that $M$ and $y \in D$ relative to model $\mathcal{M} = \langle I, D \rangle$ and assignment $g$ if and only if it is identical with a sentence $(\exists x)\neg p$ and sentence $(\exists x)p$ carries ontological commitment to entities $\alpha$ such that $M$ and $\alpha \in D$ relative to model $\mathcal{M} = \langle I, D \rangle$ and $g$.

Next follow sentences where a conjunction is in the scope of the objectual existential quantifier.

**Metadefinition 22** Sentence $\phi$ carries explicit ontological commitment to entities $\alpha$ such that $M$ and $N$ and $\alpha \in D$ relative to model $\mathcal{M} = \langle I, D \rangle$ and assignment $g$ if and only if it is identical with a sentence $(\exists x)(p \& q)$ and the sentence $(\exists x)p$ carries ontological commitment to entities $\alpha$ such that $M$ and $\alpha \in D$ relative to model $\mathcal{M} = \langle I, D \rangle$ and $g$ and the sentence $(\exists x)q$ carries ontological commitment to entities $\alpha$ such that $N$ and $\alpha \in D$ relative to model $\mathcal{M} = \langle I, D \rangle$ and $g$.

Next follow sentences where a disjunction is in the scope of the objectual existential quantifier.

**Metadefinition 23** Sentence $\phi$ carries explicit ontological commitment to entities $\alpha$ such that $M$ or $N$ relative to model $\mathcal{M} = \langle I, D \rangle$ and assignment $g$ if and only if it is identical with a sentence $(\exists x)(p \lor q)$ and sentence $(\exists x)p$
carries explicit ontological commitment to entities such that \( M \) and \( \alpha \in D \) relative to \( \mathcal{M} = \langle I, D \rangle \) and \( g \) and the sentence \((\exists x)q\) carries ontological commitment to entities \( \alpha \) such that \( N \) and \( \alpha \in D \) relative to \( \mathcal{M} = \langle I, D \rangle \) and \( g \).

Next we must consider sentences with more than one quantifier.

**Metadefinition 24** Sentence \( \phi \) carries explicit ontological commitment to entities \( \alpha \) such that for some entity \( \beta \) \( M(\beta/\alpha) \) and \( \alpha \in D \) relative to model \( \mathcal{M} = \langle I, D \rangle \) and assignment \( g \) if and only if it identical with a sentence \((\exists y)(\exists x)p\) and for all entities \( \beta \) the sentence \((\exists x)p\) carries explicit ontological commitment to entities \( \alpha \) such that \( M \) and \( \alpha \in D \) relative to model \( \mathcal{M} = \langle I, D \rangle \) and assignment \( g(\beta/y) \).

Formal implicit ontological commitment can be defined with the aid of explicit ontological commitment in a way similar to the way it was defined before.

**Metadefinition 25** An assertion \( \phi \) in a language \( L \) carries formal implicit ontological commitment to entities \( \alpha \) such that \( M \) and \( \alpha \in D \) relative to model \( \mathcal{M} = \langle I, D \rangle \) and assignment \( g \) if and only if there is such a statement \( \psi \) that \( \psi \) carries explicit ontological commitment to entities \( \alpha \) such that \( M \) relative to model \( \mathcal{M} = \langle I, D \rangle \) and assignment \( g \) and for all models \( \mathcal{M}' = \langle I', D' \rangle \) where \( I' \) is defined on \( L' \) and assignments \( h \) \( \phi \) is true with respect to \( \mathcal{M}' \) and \( h \) only if \( \psi \) is true with respect to \( \mathcal{M}' \) and \( h \) (i.e. \( \models_{\mathcal{M}',h} \phi \rightarrow \models_{\mathcal{M}',h} \psi \)).

Broad analytic implicit commitment can also be relativised to interpretations and models. This requires that the notion of a model is altered so that meaning postulates are taken as an element of a model; a model is then a triple of an interpretation function, a domain and a set of meaning postulates (which I symbolize as \( \mathcal{P} \)).

**Metadefinition 26** An assertion \( \phi \) in a language \( L \) carries broad analytical implicit ontological commitment to entities \( \alpha \) such that \( M \) and \( \alpha \in D \) relative to model \( \mathcal{M} = \langle I, D, \mathcal{P} \rangle \) and assignment \( g \) if and only if there is such a statement \( \psi \) that \( \psi \) carries explicit ontological commitment to entities \( \alpha \) such that \( M \) relative to model \( \mathcal{M} = \langle I, D \rangle \) and assignment \( g \) and for all models
\[ M' = \langle I', D', \mathcal{P}' \rangle \] in which the meaning postulates \( \mathcal{P} \) hold i. e. \( \mathcal{P}' = \mathcal{P} \) and where \( I' \) is defined on \( L' \) and assignments \( h \phi \) is true with respect to \( M' \) and \( h \) only if \( \psi \) is true with respect to \( M' \) and \( h \) (i. e. \( \models_{M', h} \phi \rightarrow \models_{M', h} \psi \)).

We can also form analogously to the concept of logical or formal truth, which is commonly taken to be truth relative to all models, a concept of logical or formal ontological commitment, a concept of carrying ontological commitment relative to all models (and not just relative to all domains compatible with a single interpretation). We can say that a sentence carries explicit formal ontological commitment to some entities iff it carries commitment to them relative to all models. We can also say that a sentence carries logical ontological commitment to some entities iff it carries implicit formal ontological commitment to them relative to all models, i. e. iff it implies a sentence which carries explicit formal ontological commitment to them. However, obviously very few sentences carry explicit formal ontological commitment to anything, since such sentences would have to contain no other constants than logical constants, since if they contain non-logical constants their interpretation depends on the model. Therefore even the implicit formal ontological commitment of most sentences is rather trivial, so this is not a very useful concept.

However, there are some moderately interesting examples of explicit formal ontological commitment. The sentence \((\exists x)(\exists y)(\exists z)(x \neq y \neq z)\) carries explicit formal ontological commitment to entities coexisting with at least two other entities relative to all models, since the models in which it is true are the same as the models in which at least three entities exist. Therefore also any sentence which logically implies this sentence carries logical ontological commitment to entities coexisting with at least two entities.\footnote{The question of the range of explicit formal commitment is connected to a difficult and controversial question. Just as it is controversial whether sentences with higher-order quantification are logically true, so it is controversial whether sentences with higher-order quantification are ontologically committed to entities which the predicate variables bound by existential quantifiers occurring in them range over. If existentially quantified predicate variables are also allowed to carry ontological commitment, contrary to what George Boolos has argued in \cite{Boolos84}, then the range of sentences with explicit formal ontological commitment will expand greatly. Unfortunately, I do not have the space to treat the question of the commitment of higher order quantification in this dissertation. I will only note two things. First, Quine and Church themselves thought that quantified predicate variables carried ontological commitment. Quine, of course, might have thought this only because he thought that higher-order quantification was always reducible to first-order quantification; however, it is doubtful if he would have denied that such quantification exists.}
Though we can thus form a model-theoretic conception of ontological commitment, we must not exaggerate its importance. The model-theoretic view of truth itself cannot replace Tarski’s original theory of truth, and similarly the model-theoretical conception of ontological commitment cannot replace the absolute conception of ontological commitment. It is true that model theory is in some respects more general than Tarski’s original theory of truth. However, it is not a generalisation of Tarski’s original theory of truth in the sense that Tarski’s theory of absolute truth would follow from model theory. Tarski’s original theory of truth is also in some respects more general than model theory. A theory telling what are the models which satisfy a set of statements does not imply any theory which would tell what are the truth-conditions of the statements contained in that set. Even if we know that “Snow is white” is true in those models whose interpretation function associates “snow” with some object that falls into the class with which the interpretation function associates the word “white”, this does not tell us that it is true if and only if snow is white. Models whose interpretation function associates “snow” with water and “white” with wet substances and in whose universe there is water would also make the sentence true. Model theory cannot by itself distinguish these two kinds of models and hence cannot give the truth-conditions of any sentence in any interpreted language, whether artificial or natural. However, the knowledge of truth-conditions is what we need if we are to use semantics for ontological purposes. Thus while model theory can be useful if used to supplement Tarski’s original theory of truth, as it offers some extra generality, it cannot replace it. Similarly the model-theoretical conception of ontological commitment cannot replace the absolute conception of ontological commitment.

\[\text{carried ontological commitment even if he would have rejected this reducibility thesis.}
\text{In any case, this is not the reason for Church’s opinion, since Church never maintained}
\text{any such reducibility thesis. Secondly, Boolos allows that sentences with quantified predicate variables have truth-conditions but does not admit that they would have ontological commitments. This goes against the parallelism between truth-conditions and ontological commitment I have stressed in this dissertation, which suggests that his views may not be fit together with those expressed in this dissertation. Here the intuitions of different philosophers conflict, and I would rather trust Quine and Church than Boolos.}\]
Chapter 6

What is the Problem of Universals?

6.1 The Problem of Universals, Abduction and Induction

In this section of this dissertation I will apply the results of the former sections to examining a metaontological problem, which has lately been thematized for example in [RP00]: what is the problem of universals? I will argue for a different solution from the one that has recently been popular, a solution that represents a return to an older view.

What is the problem to which the postulation of universals or alternatively of tropes (or of both in baroque ontologies like that of Barry Smith) is a solution, and what kind of method could be used to defend an answer to it? A very popular position, recently defended by for example Gonzalo Rodriguez-Pereyra in [RP00] and in [RP02] and by Chris Swoyer in [Swo96, page 247] is that the problem could only be the problem of how something is possible and the method must hence be inference to the best explanation, i. e. abduction in the sense in which Peirce used the word. Abduction is often explicitly contrasted with conceptual analysis, which is held to be a wrong way to solve the problem. In this dissertation I will give an argument against this popular supposition, arguing that the problem is instead just a problem of what exists, and that in order to solve its main part demonstrative reasoning is enough, though abductive reasoning or some other form of non-demonstrative reasoning may also play a part in solving other parts of
I will concentrate mostly on Rodriguez-Pereyra’s version of this supposition, since he makes the most use of it and presents it most clearly. It is vital for his argument for his theory of universals, a version of resemblance nominalism. Thus if he can be shown to be wrong in this claim, then most (though perhaps not all) of his arguments for resemblance nominalism will collapse. However, I will also examine briefly what use other philosophers have made of this claim and how they argue for it, since my ultimate aim is to show that the assumptions common to all of them are false.

The importance of abduction has long been a matter of dispute in the philosophy of science as well as general epistemology. The dispute has continued in the modern philosophy of science, and though abduction is the fashionable choice, there yet remain many champions of induction. The metaontological view that abduction is central in metaphysics is clearly a generalization into metaphysics of the view that it is central in sciences, so here the problems of metaontology and philosophy of science are closely connected. A naturalistic approach to metaphysics would say that the solution to the metaontological problem would depend entirely on the solution to the problem in the philosophy of science; however, even extreme naturalists can differ in their views of the correct method of science and hence in their views on the correct method in metaphysics. Even a philosopher who is not a naturalist would have to say that the metaontological problem cannot be solved without a thorough review of the problem in the philosophy of science.

Some philosophers (such as van Fraassen in [VF89]) have attacked the very existence of valid inference to the best explanation, claiming that a theory can never be justified just by showing that it would explain some-
thing. Even philosophers who defend abduction as a method hold two views of abduction; some abductivists hold abduction to be a method of justification, but others hold it to be just a method of discovery. Sami Paavola in a study of abduction [Paa06] makes a distinction between Harmanian abduction (abduction as a method of justifying a theory) as presented by Gilbert H. Harman in [Har65] and Hansonian abduction (abduction as a method of discovery) as presented by Norwood Russell Hanson in [Han58]. According to Paavola, Peirce himself changed his views about what abduction was, at first claiming it to be a method of justification but later thinking it to be just a method of discovery. The views of Rodriguez-Pereyra and Swoyer apparently demand the validity of Harmanian abduction. Of course if abduction is not a method of justification but at most of discovery, then an ontological theory such as a solution to the problem of universals could not be justified abductively either.

Harman himself admitted that there were unsolved problems in the theory of abduction. He said in [Har65, page 89]:

There is, of course, a problem about how one is to judge that one hypothesis is sufficiently better than another hypothesis. Presumably such a judgment will be based on considerations such as which hypothesis is simpler, which is more plausible, which explains more, which is less ad hoc, and so forth. I do not wish to deny that there is a problem about explaining the exact nature of these considerations; I will not, however, say anything more about this problem.

Unfortunately, though the theory of inference to the best explanation has been defended by many philosophers after Harman, none of them has said much about this problem either. Because of this the theory of inference to the best explanation is at present more a sketch of a theory rather than a fully developed theory. However, there exist theories of induction developed in detail and rigorously on the basis of the work of such philosophers as Carnap (for example in [Car71]), even though serious problems are known to exist in all of them. Because of this the theory of induction seems to me to be in a far better shape than the theory of Harmanian abduction as an account of non-demonstrative reasoning, though no one is in a position to say certainly which is closer to truth.
With van Fraassen his attack on abduction is combined with constructive empiricism, and is a part of his attack on scientific realism, the view that the entities postulated by scientific theories (including unobservable entities) really exist. Such views may have caused as a reaction to them abductivism to be popular among those philosophers of science that are in favour of scientific realism. However, there is no reason why an opponent of abduction (as a primitive method of justification) would also have to be an opponent of scientific realism. It is quite consistent and reasonable for a scientific realist to claim that we gain our knowledge of unobservable entities with the aid of induction instead of abduction. Actually van Fraassen rejected induction as well as abduction. His own preferred method of revising empirical beliefs was Bayesian conditionalization; however, in a Bayesian theory a priori probabilities are allowed to be arbitrary and it is surely possible to support beliefs regarding unobservable entities by Bayesian conditionalization if the a priori probabilities are chosen suitably.

As Nelson Goodman showed in [Goo54], the problem of induction is connected with the question of what predicates (or properties) are projectible and why, so whether we can inductively infer the existence of unobservable entities depends on the connection between the notions of projectibility and observability. I will examine the concept of projectibility and its possible ontological implications at more length later in 6.3.3; those readers who are unfamiliar with the notion might want to skip ahead and check its description there. Here I will only note that in order to claim that induction can reveal unobservable properties for us one must just hold that a predicate or property can be projectible even if some entities belonging to its extension are unobservable (i. e. if the predicate is weakly theoretical in the sense I have defined before in Section 3.3). However, this is a quite plausible claim. There may be indeed some reasons (even if not very strong ones) to think that projectible properties must be weakly observable (as I will argue later in 6.3.3) and therefore some of their instances must be observable. However, there is surely no good reason to think that all instances of projectible properties would have be observable, i. e. no good reason to think that projectible properties would have to be strongly observational. Indeed, Goodman himself held that they did not have to be, saying in [Goo66, page 330] that

the line between observation-predicates and others does not co-
incide with the line between projectible and nonprojectible predicates.

Goodman gives as an example the predicate "conducts electricity" which he apparently thinks is projectible but not observable.

Harman claimed in [Har65, page 90] that the inference from experimental data to the theory of subatomic particles certainly does not seem to be describable as an instance of enumerative induction. He does not make clear what the difficulty he seems to perceive is; of course it would be extremely difficult (and perhaps impossible in practice) to show in detail how an extremely complex theory such as modern atomic theory can be justified inductively, but of course it would also be extremely difficult (and perhaps impossible in practice) to show in detail how it could be justified abductively, for that matter, even if it could be. However, contrary to what at least van Fraassen if not Harman would claim, at least the unobservability of subatomic particles does not lead to any difficulty of principle in justifying their existence inductively. If we hold along with Goodman that predicates can be projectible even if some entities to which they apply are not observable, we can easily explain in broad outline how we can arrive inductively at the existence of unobservable entities such as subatomic particles. Let us assume that the predicate ",Q" is a monadic predicate that can be applied to bodies of any size and R a relational predicate that can hold between objects of all sizes; clearly there are many such predicates such as (in the case of the monadic predicate) predicates indicating the mass or charge or shape of bodies, and more complex predicates constructed out of these. If all bodies we have observed that are Q have such parts x₁, x₂, . . . and xₙ that R(x₁, . . . , xₙ), we can in many cases infer inductively that any body that is Q (probably) has such parts x₁, x₂, . . . and xₙ that R(x₁, . . . , xₙ). More exactly, we can make such an inference in all cases where the complex predicate "has such parts x₁, x₂, . . . and xₙ that R(x₁, . . . , xₙ)" is projectible (with respect to the predicate "Q"). Intuitively many complex predicates like this are projectible². If this predicate is projectible then even if some

²The greatest problem in the theory of induction is finding out what predicates (or properties) are projectible. This problem was discovered by Goodman. Goodman’s own solution was that it was predicates that were entrenched, i. e. used already frequently in sciences. I do not think that this is any genuine answer to the problem; no doubt most predicates long used in sciences are projectible, since scientists tend to use projectible predicates, as they have an intuitive sense of what predicates are projectible, but I do not
body which is $Q$ is a minimal observable body, so that we cannot observe whether it has parts, we can still inductively infer that it (probably) has parts between which the same relationship $R$ holds. In this way we can get inductive knowledge of bodies which are too small to observe.

Therefore I think that the worries of philosophers like van Fraassen about the validity of Harmanian abductive reasoning must be taken seriously even by those who are not convinced by other aspects of his philosophy such as constructive empiricism. Nevertheless, I am not completely convinced by this line of thought, since there are cases where it is intuitively highly plausible to claim that abductive reasoning has led to results which are justified (at least to some extent), even though it is rather hard to say exactly what would make such reasoning justified in those cases. It seems likely to me that if abduction justifies some hypothesis in these cases, it is because abduction can in such cases be reduced to induction. If we think after Keynes and Carnap that induction can be based on probability, then we can answer the question that Harman leaves unanswered in outline so that an explanatory hypothesis would be better than another if it is more probable relative to the data that are to be explained.

However, I cannot justify this view in this dissertation, so I will try not to commit myself to it in the discussions that follow. I will instead argue that even if inference to the best explanation is the only correct way to solve some problems, whether in natural science or philosophy, it is not a correct way to solve the problem of universals (at least not to solve all of its parts), since the main part of the problem of universals is very different from all those problems which it is plausible to claim have been successfully solved by abductive reasoning. For instance, it is usually agreed that entities whose existence is the conclusion of abductive reasoning are unobserved (like the subatomic particles Harman referred to), but many philosophers have claimed that universals or tropes can be observed. I will argue that the only way to maintain the thesis that the problem of universals is solved by

\[\text{think that there is any reason to think that all of them would be such, since scientists can make mistakes in this matter. In any case the question remains what is common to all such predicates and how this intuition works when it does. Furthermore, this answer leads to a too conservative theory of science; in any great scientific revolution not only new hypotheses are formed but also new concepts are formed, which are essential to such new hypotheses, and unless some of them could be considered projectible, few new theories could be regarded as better than old ones. Goodman’s theory would stifle all progress in sciences.}\]
inference to the best explanation is to admit that the word “explanation” is used here in a more general sense than in the philosophy of science. However, in this case inference to the best explanation can no longer be contrasted with conceptual analysis, since conceptual analysis turns out to be one form of explanation in this very wide sense of the word, so the popular position is in any case wrong in resisting the use of conceptual analysis in the solution of the problem of universals.

Rodriguez-Pereyra argues that since the problem is an explanatory one a solution to it must account for the truthmakers of certain truths; I will argue that since the problem is not (at least not solely) an explanatory one but (at least in part) a descriptive one a solution to it must account for the ontological commitments of some truths also (or even instead of truthmakers). I will then suggest a preliminary solution to this problem of ontological commitments.

It is even contentious whether the problem of universals is an ontological problem at all, and hence whether we are here engaging in metaontological research at all. Many philosophers have claimed that the problem of universals is really a problem in the philosophy of language or more specifically semantics or possibly in epistemology or more generally in phenomenology. For instance, Gregory Salmieri thinks in [Sal08] that the problem is more accurately called the Problem of Concepts, and says that the problem is how unitary thoughts that apply to a plurality of objects are to be explained and by what standards they are to be evaluated. However, I will argue that while there are problems in the philosophy of language and other such fields closely related to the problem of universals, so that it would also make sense to call one of them a problem of universals, yet there is clearly also an ontological problem which deserves that name. Indeed this ontological problem has historically been called the Problem of Universals first and perhaps also most commonly and therefore deserves that name most of all these problems, though here it is obviously not a question of absolute correctness but just of the appropriateness of nomenclature, so there is some room for stipulation.

I would then say that the Problem of Concepts that Salmieri formulated is a quite legitimate problem and is very closely connected to the Problem of Universals, but it is yet not the same problem. Indeed, the Problem of Concepts may even be epistemically a more fundamental problem than
the Problem of Universals, but it is also a more extensive problem, so it is not reasonable to try to solve it entirely before solving the Problem of Universals. It is not the same and it is a more extensive problem because a complete answer to the Problem of Universals might leave open many parts of a complete answer to the Problem of Concepts. An ontologist is only interested in an epistemological or semantical problem such as the Problem of Concepts so far as it affects ontological questions. This means that in this case he is only interested in finding out what entities we have to suppose there are in order to account for the application of concepts to the plurality of objects. However, answering that problem need not give a complete explanation of how concepts apply to a plurality of objects.

6.2 A Problem of Explanation or of Description?

According to many ontologists such as those I have already mentioned universals (or tropes or both) should be posited (if they should be posited at all) to explain various phenomena, such as how general terms can apply to different individuals or even how different individuals can be of the same kind or have the same property. That the problem would be of this kind was already suggested by Armstrong in [Arm78, §4 vii, page 14] when he said

The problem of universals is the problem of how numerically different particulars can nevertheless be identical in nature, all be of the same type.

However, with Armstrong this was just an incidental remark, and as I will show later in 6.4.1, there are also in Armstrong suggestions of a different theory. Many recent ontologists, e. g. Rodriguez-Pereyra in [RP00] and Chris Swoyer in [Swo96, page 247], make this claim central to their theory and draw strong conclusions from it. Even philosophers to whose thinking it does not really fit (as will show later) have adopted this formulation of the problem, for example Rosenkranz in [Ros93, page 72]

Rodriguez-Pereyra claims in [RP00] that the problem of universals is

the problem of showing how numerically different particulars can have the same properties.
Rodriguez-Pereyra refers to Nozick [Noz81, pages 8-9], according to whom many philosophical problems have the form: how is a certain thing possible given (or supposing) certain other things? According to Rodriguez-Pereyra [RP02, §1.1, page 19] the problem of universals is also of this form, which is seen most clearly in Armstrong’s formulations of the problem.

It is not at first sight clear just what such a question as what a philosophical problem is amounts to. Any such question can be understood either as a purely descriptive, historical question of how the problem has in fact been formulated or as a partly normative question of how it should be formulated. I want to argue that the problem is in part historical and in part normative. There is no point in calling a problem a problem of universals unless it has some connection with the historical problem; however, this might leave a lot of leeway about how to formulate the problem, especially if the historical formulations turn out to be ambiguous or if different historical formulations turn out not to be equivalent with each other, since in such a case we will have to ask which reading or formulation is more important and fruitful. However, I will argue that there are reasons to doubt both the historical accuracy of Rodriguez-Pereyra’s view as a description of how the problem has in fact been formulated and (of course more importantly) the fruitfulness of this view as a normative prescription of how the problem should be formulated. I will discuss the reasons for these two doubts in this order.

6.3 The Problem historically considered

I admit that Armstrong’s formulation, which I have already quoted, does indeed clearly indicate that the problem would be of the form of the problems Nozick refers to. However, the problem is that almost no formulation of the problem other than those of Armstrong or those influenced by Armstrong makes the problem seem similar to Nozick’s problems. This creates a serious problem with Rodriguez-Pereyra’s claims about the nature of the problem of universals. He claims that the problem he is dealing with is very old - according to him it has been discussed at length for many centuries and (see [RP02, page 1]) was already discussed by Plato. However, despite this Rodriguez-Pereyra bases his claim about the nature of the problem on formulations all of which are all rather recent, given by modern philosophers, especially on the formulation given by Armstrong, who certainly was not
a contemporary of Plato and did not live centuries ago. It might be that
their formulations are not general enough to capture what is common to
all who have treated the problem through the centuries, but at most only
what is common to (most of) those that have treated the problem lately. In
that case it might be that they do not capture the problem itself in its full
generality but only a subproblem of it on which attention has been focused
in modern times. If we look at the historical evidence, we see that this is
indeed the case. It appears that Rodriguez-Pereyra has confused a centuries
old problem with a problem which is not quite thirty years old!

6.3.1 Porphyry’s Formulation of the Problem of Universals

The problem of universals was famously first posed by the Neoplatonist
philosopher Porphyry in an introduction to Aristotle’s *Categories*. Porphyry
formulated the problem in this work but did not himself offer a solution,
at least not in that same work; he had, however, another commentary on
*Categories*, this one in dialogue form and there are other works of Porphyry
where more constructive treatments to the problem have been located.\(^3\)

Porphyry was probably the first Neoplatonic philosopher who mainta-
ined the controversial claim (accepted by most later Neoplatonists) that
Plato’s and Aristotle’s metaphysical doctrines were compatible. It is gener-
ally agreed by modern commentators (with some exceptions, of course) that
this view was historically implausible; Aristotle surely seems to be arguing
fiercely against Plato, though it is not easy to formulate exactly in termi-
nology understandable to modern philosophers what the difference in their
theories was. If then Plato’s and Aristotle’s metaphysics can retrospectively
be viewed as answers to the problem of universals, then it is no wonder
that Porphyry had a problem, as he was trying to reconcile incompatible
views. It is not wholly clear what Porphyry’s own solution to the problem

\(^3\)There is much controversy concerning what the subject matter of *Categories* is,
whether it concerns metaphysics or semantics or even syntax. Some commentators such
as Walter E. Wehrle in \[Weh00\] or Lloyd P. Gerson in \[Ger\] have argued that *Categories*
was a work on semantics, not metaphysics, and this view might be used to argue that
since Porphyry formulated the Problem of Universals in an introduction to this work
(and himself held it to be a work that concerned semantics) the Problem of Universals
is not a genuinely metaphysical problem at all. However, since Porphyry after formulat-
ing the problem of universals explicitly excluded it from further, deeper consideration in
that work, this rather provides evidence that even if the Categories were concerned with
semantics (which is very controversial), the problem he formulated is not a semantical
one.
of universals was. It is likely that he had a solution. Some modern commentators contemptuously assume that he did not give a solution in the work because he did not have one, and he did not have one because he in some way misunderstood Aristotle (though given that modern commentators disagree even more about the interpretation of Aristotle than ancient ones, many of them must misunderstand Aristotle very badly, so modern historians of philosophy are scarcely in any position to be contemptuous of Porphyry). However, Porphyry did not present the problem of universals as something that would puzzle him but as something whose discussion had to be passed by in the specific work in question since it was in his view too deep for elementary students, who were its target audience. However, it is likely that as he was trying to reconcile the incompatible metaphysical views of Plato and Aristotle, he distorted both of them, so that even if he had a theory, his theory differed from both of their theories (if they had any theories at all with respect to this problem of universals), even though he himself maintained that it was identical with both of them. Naturally the likelihood that the product of such an attempted reconciliation was incoherent (even if the theories he tried to reconcile were coherent, which is itself unlikely) or at least unnecessarily complex and convoluted, is of course very great. Of course this does not prove conclusively that it was incoherent; we would need to study his works at length to be certain of this. However, this makes it antecedently unlikely that we would gain any great light on the question from such deeper study of his works, so I will not consider his opinions more thoroughly here.

The problem passed into Medieval scholastic philosophy because of the late Roman philosopher Anicius Manlius Severinus Boethius’s translation of Porphyry’s work. Boethius apparently (e. g. according to [Twe84]) proposed a solution that he derived from the Aristotelian commentator Alexander of Aphrodisias. This solution influenced later philosophers both in the Muslim world (e. g. Ibn Sīnā (980-1037), called Avicenna in the west) and such famous western medieval philosophers as the most famous scholastic philosopher, Thomas Aquinas, both through Avicenna and through Boethius’s western successors. It became the standard solution to the problem. It is a solution that is usually called Moderate Realism; however, it is so very moderate a form of realism that it is highly doubtful whether it is genuinely realistic at all. Even worse, it is also extremely obscure so that different
modern commentators on it disagree about what it really means and it can with good reason be questioned whether it is at all a coherent solution.

This theory is so very extraordinarily subtle that it is impossible to succinctly and satisfactorily summarize the theory in any generally acceptable way. However, I will try by saying that it very roughly says that the genera and species (or common natures) are universal in the mind but particulars in the sensible objects and are neither universal nor particular in themselves. The real objective foundation of universality lies according to it in the similar natures of the members of the same class. This theory is in my opinion not realistic at all, but so far as any intelligible sense can be made of it, it is a combination of concept nominalism and trope nominalism and resemblance nominalism.

The theories of earlier philosophers like Plato or Aristotle or the Stoics can retrospectively be considered as answers to the problem of universals. However, there are serious problems that are not usually seen about how to

---

4It may be true that as Salmieri says in [Sal10, page 16]
this is the position that until quite recently was held by the overwhelming majority of people who called themselves moderate realists. In particular it was the position of the neo-Scholastics from whom we have inherited the term.

I am not sure that we have inherited the term "universal" from the neo-Scholastics rather than by other ways from medieval and directly from ancient philosophy. In any case, that the adherents of this position call it realistic does nothing to show that it would be genuinely realistic, any more than e. g. the fact that Berkeley called his immaterialism commonsensical did anything to show that it would have been genuinely commonsensical. In fact while the position of Alexander was yet extremely ambiguous the position of the neo-Scholastics to whom Salmieri refers - e. g. Maurice de Wulf in [dW52] - was quite unambiguously not realistic. Salmieri expresses this view in [Sal10, page 16] so that all particulars falling under a universal are composed of two components, of which one is identical to a corresponding component in others, but this identity need not be numerical identity but can be qualitative identity. Here Salmieri understands by qualitative identity (which most modern philosophers would not call identity at all) the relationship between two things whereby there is no difference between them except in number. Qualitative identity seems to be the same as exact resemblance and the qualitatively identical components seem exactly what trope theorists call tropes, so this view collapses into a combination of trope nominalism and resemblance nominalism (with some concept nominalism on top of it according to other places of Salmieri’s article) instead of being any kind of realism at all. Salmieri very surprisingly denies that Aristotle would be a realist even in this attenuated sense. A similar position with regard to other entities would generally not be called realistic with regard to those entities, so it is singularly inappropriate to call such a position with regard to universals realistic. We would not, for example, usually call a phenomenalistic theory according to which physical objects have unity only in our minds but have a real objective foundation in the relations between (classes of) sense-data (which according to the theory exist independently of being sensed) realist with respect to physical objects.
do so. It is usually claimed in histories of philosophy that Plato and Aristotle proposed opposing answers to the problem; Plato an extreme form of realism, universalia ante rem, according to which universals existed separately from particulars and were independent of them and Aristotle a moderate form of realism, universalia in re, according to which universals existed in particulars (whatever this separation and lack of separation exactly mean). However, this is by no means as clear as it is often supposed to be; there is no agreement among historians of ancient philosophy whether Aristotle was a moderate realist or not a realist at all. It is unquestionable that Aristotle’s philosophy was realistic with respect to many questions (at least he was realistic about biological organisms and the heavenly bodies, though less so about such physical objects as rocks) but it is far from clear that he was realistic about universals. Gregory Salmieri argues in [Sal10] that in fact it was Eudoxus of Cnidus of the early group of followers of Plato called the Old Academy that proposed a theory of moderate realism, of universalia in re, as an alternative to Plato’s extreme realism. It would then be the little-known, almost forgotten Eudoxus that is the true originator of the kind of moderate realism today defended by ontologists like Armstrong, not Aristotle. According to Salmieri Aristotle attacked this moderate realism as well as Plato’s original theory, and Aristotle’s own theory was far more anti-realistic or deflationary concerning universals, in fact a form of conceptualism.

However, the standard story can be questioned even more radically; it may not be just that Aristotle proposed a different answer to the problem of universals than superficial popular histories of philosophy commonly claim but that he proposed no answer at all. Plato and Aristotle certainly did not propose such competing theories explicitly, since they were not explicitly aware of the problem. Plato may not have had any concept of universals; as we will later see, it was probably Aristotle who first coined that concept. Aristotle argues against the view that universals are substances, and seems to take this to be Plato’s position (as did many later Aristotelians), but as e. g. Lloyd P. Gerson shows in [Ger], it is far from clear that Plato did this, especially as it is unlikely that Plato already possessed either Aristotle’s concept of a universal or any of Aristotle’s concepts of a substance. Plato nowhere says that the forms are predicated of sensible particulars, and it is unclear if the participation that he speaks of can be identified as
metaphysical predication i.e. instantiation or exemplification. Aristotle had the concept of a universal, which he may have invented himself, but it is not clear that he ever asked himself any question similar to the one Porphyry asked. He may have given an implicit answer to that question, but commentators have difficulty agreeing on what it might be. It may even be that the disagreement results from asking a question with no answer. For example, Riiu Sirkel claims in [Sir10, page iii] that Aristotle’s positive remarks about universals remain neutral with regard to their ontological status. Paradoxical as it may sound, it might even be that Plato’s and Aristotle’s metaphysical theories are irrelevant to the problem of universals. However, more likely they do have some relevance to it (and Sirkel also in the end finds out an implicit position about the ontological status of universals in Aristotle), but finding out what kind of relevance they have is harder than is commonly thought. However, happily solving this historical question is not very important since solving it would not be likely to help us much in finding a solution to the problem of universals that would satisfy us rationally.

Though the problem was then implicitly addressed already by Plato and Aristotle, it was Porphyry who first explicitly thematized the problem. As

---

One of the most important reasons to think that Plato’s theory of Forms or Ideas and Aristotle’s hylomorphic metaphysics are relevant to the problem of universals is that one of the Greek words translated as “form” (besides μορφή, morphee) is еιδος, eidos, the same word that Porphyry uses and that is translated as ”species” in his formulation of the problem of universals. A problem, however, is whether the word is used in the same sense or in different but equivalent senses or on the contrary in non-equivalent senses when speaking of forms and speaking of species. Is there indeed any reason to use two words in translating Aristotle’s eidos, both ”form” and ”species”, or could one of these be discarded entirely? Could we for instance say that according to Aristotle a composite sensible substance is composed of species and matter? Commentators on Aristotle do not seem to often notice this question and just suppose without any justification that the Greek word has two meanings, perhaps just because this is how it has been traditionally done. Owens notes in [Owe63, page 159] that commentators such as Hermann Bonitz have noted that for Aristotle these two meanings (one of the four causes and the subdivision of a genus) somehow coincide. If they coincide why cannot they just be identified? However, the fact that Aristotle seems to say in Metaphysics that forms are primary substances (1037a21-1037b7) and also that universals are not substances (1038b34-1039a2) suggests that at least in Aristotle forms cannot be the same thing as species or even coincide with them, since Aristotle clearly assumes that species are universals. Of course, these statements might come from different periods of Aristotle’s development as commentators like Werner Jaeger would think, and in this case forms could after all be the same as species according to Aristotle at some time in his development, so the historical question of the relevance of Plato’s and Aristotle’s metaphysics of forms to the problem of universals is very difficult indeed.
Rodriguez-Pereyra says that he speaks about the traditional problem of universals he should be speaking about the problem Porphyry formulated. However, when we look closely at Porphyry’s text, we see that when Porphyry posed the problem, he did not formulate it as a problem of how something can be the case, but as a problem of whether something is the case. Porphyry asks (see [War75])

1. whether genera or species exist in themselves or reside in mere concepts alone;

2. whether if they exist they are corporeal or incorporeal and

3. whether they exist apart or in sense objects and in dependence

Rodriguez-Pereyra also identifies the problem he is trying to explicate as the problem of the One over Many. However, it is doubtful whether any such problem is a traditional problem; no problem with such a name has been discussed for many centuries as Rodriguez-Pereyra claims the problem of universals has been. Such a problem is only mentioned (perhaps first by Devitt in his criticism of Armstrong in [DS87, 14.2]) in recent literature dealing with the problem of universals, and is thus a recently coined problem, perhaps indeed a problem first discovered by Armstrong and named by Devitt. In older discussions of the problem of universals and recent commentaries on them philosophers usually speak about the One over Many argument: see for example [Fin80]. This meant originally an argument apparently used by Plato for the existence of ideas corresponding to all general terms (a view that can be viewed as a crude predecessor of promiscuous realism). This argument has (like a lot of arguments in the history of ancient philosophy) a rather complicated history; our knowledge of it comes to us at third hand. The argument was attacked by Aristotle in his lost work Peri Ideon which ancient commentators on Aristotle referred to. Aristotle identified the argument as such (according to the commentators) and there is some controversy over whether Plato really had such an argument or if he had it whether its conclusion was that ideas corresponded to every general term. However, the phrase has also been used to refer to more modern arguments for the existence of universals - in a sense of the word in which universals are not quite the same thing as Plato’s ideas - which have some significant similarities to Plato’s argument as formulated by Aristotle. As these are arguments for the existence of something - ideas or universals - they presuppose a problem that is a problem of whether something exists, not a problem of how something can be the case. A one over many principle is also often mentioned; it is a principle used in a one over many argument. However, naturally once you have an argument that something exists, there emerges the problem of whether that argument is valid and if not what is wrong with it, i. e. how the apparent validity of the argument can be explained, and that seems to be the One over Many problem Devitt talks about. It is thus really a subproblem of the traditional problem of universals, in the sense that solving it will help to solve the traditional problem but is not in itself sufficient and perhaps not even necessary for solving the traditional problem.

According to another translation in [Sor04, 5(4), page 157] Porphyry asks

a. whether genera and species are real or are situated in bare thoughts alone,

b. whether as real they are bodies or incorporeals and

c. whether they are separated or in sensibles and have their reality in connection with them.

These two translations do not seem to differ in any philosophically significant way.
Thus a problem of how something can be the case is not the same problem with which the medieval philosophers struggled and whose discussion has continued until the present day. Nor does the problem as Porphyry formulated it have at all the form of Nozick’s problems; it is not a problem of explaining anything at all.

However, finding that it is not an explanatory problem does not yet make clear exactly what the positive character of Porphyry’s problem is. The formulations of all three questions can be interpreted in different ways. Even if we consider only the first and most fundamental of them, we find room for ambiguity. What does Porphyry’s contrast between existing in themselves and residing in concepts really come to?

For something to reside in mere concepts could mean that it does not really exist but is only thought of, like imaginary entities such as fictional characters, or it could mean that it actually exists but is somehow mind-dependent.

E. g. if there are individual accidents of the human mind that depend on the human mind with rigid existential dependence, as Aristotelians famously thought, then these accidents are mind-dependent without being imaginary, and many medieval nominalists or conceptualists identified universals with such accidents of the mind, such accidents as would be called concepts. A more subtle version of such a theory would not identify universals with anything, but would take them to be logical constructions out of such individual accidents of the mind.

However, a nominalist could also think that universals are imaginary, just like unicorns or witches are usually thought to be, i. e. in phenomenological terms be purely intentional objects having only intentional inexistence (or in Roman Ingarden’s terminology existentially heteronomous objects)⁸. One way of developing such a theory would be to propose a fictionalist theory of discourse concerning universals. The ontological status of universals would

⁸This corresponds also to Jody Azzouni unusually strong notion of ontologically dependent objects (as explained in [Azz04, page 96]); entities which would be ontologically dependent in a sense of the word more commonly used by ontologists, such as Aristotelian accidents, would not be ontologically dependent in Azzouni’s sense. Azzouni thinks that entities which are values of variables bound by an objectual quantifier could still be ontologically dependent in this strong sense; however, I have already given arguments for the contrary view that any objects which are values of a variable bound by an objectual quantifier must already be ontologically independent in Azzouni’s sense of the word.
then be similar to that of such particulars as Donald Duck.9

There may also be other ways to understand Porphyry’s contrast, though I cannot think of any even prima facie plausible ones.

Happily this question does not matter much. The problem of universals can be understood as dispute between realism and anti-realism with respect to certain specific purported entities, namely universals. A genuine realism concerning universals, like genuine realism concerning any entities, requires both that they exist and are independent of the human mind and language, so if we are to make a decision between realism and nominalism we must ask both questions. Also it is obviously pointless to ask whether something is mind-dependent or language-dependent or independent of mind and language if we do not know whether it exists at all; thus the first question must be whether there are universals at all. An eliminativist resist physicalist might hold that there would not be universals even in a non-realistic sense.

A historical perspective is thus useful, though we must not tie the concept of universals dogmatically to any historical theory of universals if it is to be a concept that is to be of any relevance to present day discussions. I must stress that the purpose of this dissertation is not to be a part of history 9

---

9These two kinds of theories have both in fact been held by nominalists throughout the ages. As expounded for example in [Rea77, page 15], William of Ockham changed his mind from the latter of these theories to the former. At first he thought (influenced by Henry of Harclay) that a universal was a fictum, having merely objective being (by which the scholastic meant something almost opposite to what the term would mean today, namely the same as the intentional inexistence of the phenomenologists). However, later (under the influence of Water Chatton) he thought that universals existed subjectively, as qualities of the mind or mental acts. A translation of a passage where William debates the relative merits of these two nominalistic theories can be found in [Spa94, pages 218-230].

10Porphyry did not explicitly consider the possible theory that universals might be dependent on language, since he thought as most ancient philosophers had done that linguistic entities reflected mental acts (spoken words were symbols of affections or impressions of the soul as Aristotle says at the very beginning of De Interpretatione 16a4-5), and therefore if universals were to be reduced to general terms they would be further reduced to the general concepts of which such terms were mere symbols. However, the theory that universals depend on language (which is called nominalism in a strong sense, Armstrong’s predicate nominalism) was apparently already suggested in the Middle Ages by Roscelin, who is said to have said that a universal is an emission of sound (flatus vocis), though it is not sure if this was an accurate description of his position. In any case it never became popular during medieval times, while conceptualistic theories (which can be called nominalistic in a weaker sense) which held universals to be in some sense mind-dependent became most popular in later scholasticism. However, nominalism in the strong sense has become a popular position in recent philosophy so it cannot be left out of account. In any case mentalistic and linguistic subjectivism are in many ways parallel and come up in later philosophy when the question of realism is asked with respect to any domain of purported entities, so they are best treated together.
of philosophy but an original, systematic discussion of metaontological and ontological questions. As an auxiliary for such systematic discussion I am here considering the way the problem of a universal has been formulated in early history rather than arguments that have then been given for or against any solution to it. These arguments are often based on outdated premises such as outdated natural science (physics, biology etc.) and logical and theological assumptions that many philosophers do not take seriously today. For example, as we can see in [Spa94, page 119] Ockham argues against realism about universals by claiming that there cannot be universals since

\begin{quote}
It is proper to the divine essence alone that without any division or multiplication it is really in several distinct supposita.
\end{quote}

Many other medieval philosophers argued similarly that if there were universals they would in some way trespass on divine prerogatives. Such arguments are unlikely to impress atheists, agnostics or deists such as many if not most modern philosophers are; e. g. a modern defender of universals such as Armstrong would be unlikely to be impressed by the claim that universals would poach on divine prerogatives, as he is an atheist. Also since this argument of Ockham appeals to the doctrine of the Holy Trinity, it is not likely to impress even theists if they are not Christians\footnote{There are of course theistic arguments against strong realism about universals that do not depend on trinitarian doctrine; e. g. it is often claimed that strong realism leads to pantheism, but the arguments for this seem to be based on the Noe-Platonistic identification of God with Being, and should be attractive only to Neo-platonists}. Not that even Christians would have to endorse this kind of argument; many Christian philosophers such as Wolterstorff and Plantinga would say that the assumption that there are universals and even that they are independent of God is quite compatible with their Christianity, though of course there are also modern theist philosophers who think that the assumption is not compatible with Christianity. Of course, similarly unimpressive arguments were also presented in favour of universals in medieval times; e. g. Roscelin’s form of nominalism was argued to imply tritheism and so be inconsistent with monotheism. Medieval philosophers also tended to suppose that if something is shown to have been Aristotle’s view, it was thereby shown to be true unless it contradicted the Bible or the Church Fathers, and therefore they confused the question of what is the correct solution to the problem of
universals with the question of what Aristotle thought was the correct solution to the problem of universals; modern philosophers (with the exception of Catholic philosophers whose intellectual honesty is often open to doubt) have less respect for the authority of Aristotle\textsuperscript{12}.

It is, indeed, desirable to broaden the problem slightly from the one Porphyry originally formulated. This becomes obvious once we consider the interpretation of the words Porphyry uses further. What does Porphyry mean with the vitally important words "species" and "genus"? These words have been used in different ways throughout the history of philosophy, so finding out how they are used in a particular instance is necessary for the avoidance of misunderstanding.

Fortunately Porphyry gives some explanation for his terminology. Porphyry says later in [War75, pages 29,30] that the word "genus" has three senses, but discussion among philosophers concerns one according to which a genus is that predicated essentially of many things which differ in species. Similarly Porphyry says in [War75, pages 35] that the word "species" has many senses but the important sense is one in which a species is predicated essentially of many things which differ in number. These explanations of what genus and species are are not definitions, since as the translator explains, Aristotelians thought these concepts could not be defined (though this was because of a particularly narrow theory of definition, according to which all definition proceed by means of genus and differentia, and this kind of theory of definition is definitely outdated, so this should not make us lose hope of defining them). The explanations would indeed be obviously circular if they were definitions, as they determine genus with the

\textsuperscript{12}Semantical holists might deny that a problem can be formulated independently of evidence considered as relevant to it. However, such holism is obviously implausible. A Darwinist and a creationist consider the same problem concerning the origin of species and provide different and incompatible answers for it, even though they consider different kinds of evidence relevant, a creationist thinking biblical authority relevant and the Darwinist thinking it irrelevant. Similarly the modern philosopher and the medieval philosopher are considering the same question even though they differ with regard to what evidence they consider to be relevant. The problem of universals indeed differs from the question with regard to the problem of species in the respect that while the Biblical authority seems to favour the creationist answer strongly, it would not favour any answer to the problem of universals much more than another, as Biblical evidence could in the Middle ages with some plausibility be given both for and against the existence of universals. Generally, however, realism about universals seems to have been rejected for religious reasons, as the existence of universals seemed in some way to infringe on God’s prerogatives (which is of course the exact opposite of the view commonly found in popular histories of philosophy, which commonly claim that realism was supported for religious reasons).
aid of species and species with the aid of genus. Thus they leave the notions of genus and species rather obscure. They make clear, however, that Porphyry understands both species and genus as entities predicated essentially of something else. Thus the problem of universals would according to Porphyry only concern entities predicated essentially of many entities, like the substance universals, not accidentally predicated universals like redness. The essentialness of predication implies that the entities predicated of something are predicated of it necessarily, i.e. are predicated of it if it exists at all, but it also implies even more; it appears that genus and species are also sortal properties, so that they provide a principle of individuation for the entities of which they are predicated.

6.3.2 What are Universals?

Here it is useful to delve into the history of the word "universal" used in the very name of the problem of universals. While Porphyry does not speak of universals at all in his formulation of the problem, yet the species and genera of which he speaks had already been called universals by Aristotle, who was one of Porphyry’s authorities. Aristotle had actually before Porphyry given a definition for the concept of a universal that he probably introduced into philosophical terminology. This definition was such that it apparently also included entities accidentally predicated of something (and entities which do not provide a principle of individuation). Aristotle gave in his De Interpretatione 7, 17a the famous definition\(^\text{13}\) (I use here the Ackrill

\[\text{an entity that is eternal and inseparably inherent in many entities.}\]

\(^{13}\)Similar definitions have also been given independently of Aristotle and even independently of Greek influence. A similar definition was given in Indian philosophy by the Nyāya-Vaiśeṣika philosophers; who were more realistic than most Indian philosophers. According to Kisor Chakrabarti in [Cha75], the Nyāya-Vaiśeṣika philosophers defined a universal (sāmānya or jāti) as

\[\text{Inherence may probably be understood as the same thing as predication in a non-linguistic sense; even when the word “inherence” is used in discussions of Aristotle’s thoughts it is very close to the notion of predication, whether of universals or individual accidents or forms. If this inherence is the same as predication, then while the Nyāya-Vaiśeṣika definition is clearly stronger than Aristotle’s (adding additional features such as eternity to the definition of a universal) yet any entity that satisfies the Nyāya-Vaiśeṣika definition would also satisfy Aristotle’s definition. This surprising similarity provides strong evidence against relativistic theories that would claim that the problem of universals would only be meaningful in a Western framework. The Nyāya-Vaiśeṣika philosophers (or at least some of them) defended a quite robust version of realism. This realistic theory may have been more a product of Vaiśeṣika philosophers (who concentrated on ontology) than of the Nyāya philosophers (who concentrated on logic). Most other Indian schools of philosophy,} \]
translation [Ack36]):

I call universal (τὸ καθολοῦ, to katholou) that which is by its nature predicated (κατηγορηται, kategoretai) of a number of things

Aristotle apparently used the term "predication" so that it did not apply only to a relation between linguistic expressions of Greek or other natural languages, their singular terms and predicates, or a relation between linguistic predicates and things in the world, but also to a relation between entities in the world none of whose terms is a linguistic expression.

Nominalistic philosophers of course have tried to interpret Aristotle so that he would always mean with "predication" a linguistic relation (or a relation between mental concepts or a relation between mental concepts and entities). This is not a mere trivial terminological matter, as it might at first sight seem, but this interpretation of Aristotle’s terminology already leads to a nominalistic interpretation of Aristotle’s whole philosophical doctrine. However, the arguments of such nominalists often implied in their premises the truth of nominalism as an ontological theory, since they argued that Aristotle should be interpreted in this way because they thought that Aristotle’s true view had to be reasonable and that this interpretation was the only way to make what Aristotle said reasonable, i. e. to confirm to what they considered was the truth. The question of whether "predication" can stand for a non-linguistic relation is itself a part of the problem of universals, and a fair formulation of the problem should not presuppose an answer to it; realists hold that it can while most nominalists hold that it can not\(^\text{14}\)

\(^{14}\)Some modern philosophers have defined the word "universal" in a way different from the way Aristotle defined it. E. g. Rodriguez-Pereyra seems to presuppose in [RP02, §1, page 1] that something is a universal only if it can be wholly located in different places at the same time. He even calls this an Aristotelian characterization of a universal! However, in another place, in [RP02, §12.3, page 203], he says more in accordance with the traditional definition that "is a universal" can be defined as "is instantiated by something". Thus he apparently contradicts himself. Probably he presupposes that the two definitions are equivalent, so their difference does not matter; however, it has been often argued that they are not (for an argument see [Wol70, pages 231-233]). It may indeed be granted that since more than a few philosophers have used the word "universal" in this way, the word may by now have a sense in which it stands just for an entity which can be wholly located in different places at the same time. However, it is important not to confuse this sense of the word with its traditional sense, or one will commit the fallacy of equivocation.
Frank Lewis is one of the modern interpreters of Aristotle who holds that predication in Aristotle was at least in some cases wholly non-linguistic and clearly explains this position; he says in [Lew91b, page 4]

Our dominant notion of predication today is exclusively linguistic, so that both the subject and what is predicted of it are invariably linguistic items - a grammatical subject and predicate. For Aristotle, by contrast, the subject is an item in the ontology and not a linguistic item, and more often than not what is predicated is not linguistic either; it is not a predicate, but a predicatable.

Lewis calls Aristotle’s notion of predication metaphysical predication. However, there are also exactly opposite views. Lloyd P. Gerson says in [Ger, page 4] that

Predication is without exception assumed by Aristotle to be an extra-ontological category of activity.

This disagreement is, I think, the most fundamental disagreement that there exists with regard to the interpretation of Aristotle’s theory; it affects both our interpretation of Aristotle’s metaphysics and our interpretation of his logic and our understanding of the relation between the two. Clearly, either Lewis or Gerson must be wrong, and this threatens to invalidate almost everything that one of them says about Aristotle’s theory.

In fact it is not sure that Aristotle was consistent or clear in this matter. Jonathan Barnes says in [Bar07, page 120-121] the following:

Most of the numerous passage in which Aristotle discusses or alludes to subjects and predicates offer no clear answer one way or another to the question ‘What sort of item is a predicate?’ (you wouldn’t expect the texts to do so.) Several texts quite plainly indicate that Aristotle took objects, and not significant expressions to be predicates and subjects. Several texts quite plainly indicate that Aristotle took significant expressions, and not objects, to be subjects and predicates. That is to say, Aristotle was muddled or inconsistent when he thought about the status of predicates or rather (what comes in the end to the same thing), Aristotle probably never thought very long about the status of the things.
Jonathan Barnes does not link his discussion of predication systematically to the problem of universals (though he occasionally alludes to it), but since as we have shown the notion of predication is crucial to the definition of universals and the formulation of the problem of universals, what he says implies that Aristotle was also muddled or inconsistent in his view of the ontological status of universals.

However, even if Aristotle in some places took predicates to be expressions and in other places took predicate to be objects, this does not show conclusively that he was muddled or inconsistent. Aristotle may have used the word "predicate" just as he used the word "being", in analogous senses, so that the different ways in which he used the word would be instances of core-dependent homonymy; indeed, that the notion of being according to our modern view is in one sense predication makes this quite likely. Obviously metaphysical predication can be defined with the aid of linguistic predication and denotation and conversely linguistic predication can be defined with the aid of metaphysical predication and denotation, so either can be taken to be he focal meaning of the term "predication". In any case, even if Aristotle did not consciously use the notion of predication in this way, it seems most likely to me that he should have done this, i.e. that this way of developing his theory makes best sense of all of the places where Aristotle uses the notion of predication. However, if Aristotle’s theory is understood in this way, then in any case he would or should have taken the notion of metaphysical predication to be intelligible (contrary to what Gerson thinks), and this implies that the problem of universals makes sense in his framework even if he did not formulate it explicitly himself.

In any case many realist philosophers in the Middle Ages, like William of Champeaux and Walter Burley, took Aristotle’s relation of predication to

William of Champeaux may be one of those philosophers whom the history of philosophy has treated most unfairly. He is remembered only as the opponent of the famous nominalist Peter Abelard. William of Champeaux’s view is mostly known only from the attacks of his opponent, who naturally makes it sound silly. William of Champeaux’s realist theory was very similar to the kind of realist theories that are today presented by Neo-Aristotelian philosophers like Loux or Lowe, far more similar than any of the theories of more famous medieval thinkers like Aquinas, who as I have argued in an earlier footnote were not realists at all but held a position similar to a combination of trope theory and concept nominalism. One would think that this makes it historically significant and would inspire people to study it closely. However, this is not the case. While some of his writings have survived, they have not been edited much less translated into modern languages. Really, however, Abelard’s arguments are not all that impressive, at least against the realism of William, as they seem to be based on a deliberate misreading of William’s
be a non-linguistic relation (see for example [Spa96, page 144]). Therefore the word "predication" acquired such meaning then, if it did not have such a meaning before. If you remember Carnap’s distinction between semantical concepts and absolute concepts and his theory of their correspondence, then instantiation is the absolute concept corresponding to the semantical concept of predication.

Here, however, we must note a complication. Even if Aristotle is interpreted so that what is predicated in his theory is not always a linguistic item (nor a mental concept either, as older nominalists i. e. conceptualists would say), which seems a highly probable interpretation, this does not yet by any means imply that he would have been a realist about universals. What is predicated metaphysically could yet always be a trope instead of a universal, in which case all the universals would be linguistic predicates (and so themselves metaphysically tokens and particulars)\textsuperscript{16}. Nevertheless, if Aristotle were (in my view very implausibly) interpreted so that what is predicated of something in his theory is always linguistic or mental, then

\textsuperscript{16}It is just as controversial whether Aristotle thought there were tropes as whether he thought there were universals independent of human minds. Individual accidents have traditionally been ascribed to Aristotle on the basis of passages such as Categories 1a23-28 and these are very much like what are today called tropes. This interpretation was among others first presented in modern history of philosophy by J. R. Jones in [Jon49]. However, it has been challenged by many, famously by G. E. L. Owen in [Owe65]. Owen thought that Aristotle’s ontology contained only universals, and what had been thought to be individual accidents were just highly specific universals. Likewise there is controversy about the interpretation of Aristotle’s substantial forms; according to some interpretations they are particulars and seem then very like tropes; however, other interpreters view them as universals. There have been interpreters who think that Aristotle’s ontology contained both tropes and universals (e. g. Barry Smith who was inspired by this interpretation of Aristotle to develop his own baroque ontology or Michael Loux in [Lou91]), those like Owen who thought that it contained only universals, those that think it contains only tropes (which seems to have been the standard interpretation in the later Middle Ages), and those that think it contained neither. It can also be held that Aristotle at different stages of his development held different theories; for instance, Michael Frede seems to suggest in [Fre87, page 50] plausibly enough that Aristotle’s ontology in Categories contained both tropes and universals (kinds), but his later ontology in Metaphysics contained only tropes besides concrete particulars. It is fair to say that not the slightest agreement with respect to the interpretation of Aristotle’s ontology exists.
this would exclude a realistic interpretation of his theory of universals. Even when we leave historical considerations behind, we must say that the acceptance of metaphysical predication is necessary but not sufficient for realism concerning universals.

Aristotle’s notion of predication, when applied to a relation between non-linguistic entities, seems to me to be essentially the same as the notion expressed by the more modern words “instantiation” and “exemplification”. There are two differences, however. One difference is that Aristotle always thought of a property being predicated of a single individual or linguistically of the predication of a one-place predicate of one term; modern theories of predication include the predication of a relation of many arguments (or their sequence) or linguistically the predication of a many-place predicate of many terms. While Aristotle’s notion of predication is thus in one way narrower than the modern notions, in another it is wider, as it includes identity and the relation of property inclusion as a special case. In modern terms Aristotle’s ontological predication is the disjunction of monadic instantiation, inclusion and identity, since neither Aristotelian nor indeed later ancient or medieval logic clearly distinguished these three relations; any entity would be predicated of itself according to them and higher genera would be predicated of lower genera. Later Lesniewski’s ontology followed Aristotelian logic in this respect. E. g. according to the Aristotelian theory in saying that Socrates is Socrates the word “Socrates” is predicated of itself in the linguistic sense of predication and the human being Socrates is predicated of Socrates, i. e. of itself, in the non-linguistic, metaphysical sense of predication.

The philosophers like Frege and Russell who first distinguished instantiation and identity thought that the word “be” in natural languages was lexically ambiguous. This thesis has often been criticized recently. Much use has been made in this criticism of Charles C. Kahn’s discussion of the verb “be” in ancient Greek, as summarized in [Kah86]. This kind of argumentation seems often to confuse the question of what the semantics of the verb “be” (in either Greek or English) really is and the question of what the ancient Greek philosophers thought it was. Jaakko Hintikka and Simo Knuutila argue in [KH86, page x] that many philosophers and classicists have been reluctant to claim in so many words that the greatest Greek philosophers did not operate with the Frege-Russell distinction, since this would
amount to accusing them of a logical howler. Hintikka and Knuuttila seem to share this reluctance since they welcome the possibility that to acknowledge that Plato or Aristotle lacked the distinction is not to accuse them of any mistake. It shows in my view an unbecoming kind of authoritarianism in philosophers and classicists that they would be reluctant to admit that Plato or Aristotle might have made some logical mistakes. All scientists who create a discipline (as Aristotle practically created logic) make errors that are corrected in the later development of the discipline if the discipline is fortunate enough to progress after them; admitting that Aristotle made some serious mistakes does not take anything away from his incredible achievement in creating systematic logic. Claiming that Aristotle committed no howlers is close to the obscurantist claim that he perfected logic (as Kant thought) and that there was no need for modern logic. I admit of course that there are alternative logical treatments of the verb "be", and that the question of which of them is correct is far from settled; Hintikka's game-theoretical semantics might in the end provide evidence against the Frege-Russell thesis, as he argues. However, this is a question of systematic, synchronic semantics and historical studies such as Kahn's are not of any great relevance to it. Most particularly, even if it can be shown that Aristotle never distinguished between being as existence and being as predication and being as identity (which is quite possible, for though there is some apparent evidence against this, it is rather obscure) this does not even provide any good evidence that the Greek verb είναι (einai) was not ambiguous between such concepts, and naturally still less evidence that corresponding expressions in modern languages such as the English word "be" or the German word "sein" or the Finnish word "olla" would not be ambiguous in this way. While I am therefore not impressed by the criticisms, I do not have to assume that the word "be" is lexically ambiguous, since the meaning of the word can be understood to be the disjunction of instantiation and identity. However, I do have to assume that there are two very different cases of predication that it is useful to distinguish in logic and ontology even if not in natural language.

Of course, I am now leaving open whether there is any such non-linguistic relation corresponding to linguistic predication; if there is not it follows from my definition that there are no universals, so I am not presupposing the solution of any substantive ontological question here. I will try to prove
that there is such a relation in a later subsection of this dissertation.

There is also another possibility that must be taken into account; there might be many relations corresponding to linguistic predication. In this case the concept of universal would be ambiguous. In fact there are many relations associated with linguistic predication. However, at least one of these relations is such that it does not give rise to any non-trivial concept of a universal. When a predicative statement is true there exist an entity and a universal predicated of it or instantiated by it, but according to those philosophers who accept the truthmaker axiom, if the universal is not a component of the entity there exists also a fact or a trope whose existence requires that the universal is predicated of the entity. Therefore a relation would also hold between the entity and the fact or trope which implies the existence of the entity and the universal. Some philosophers, especially trope theorists such as G. F. Stout call the relation between the trope and the entity of which something is predicated in the sense I have used the word predication. E.g. if a man is happy, then while we may say that the universal happiness is predicated of him, we may also say that the fact or trope, his being happy, is predicated of him. There is also a third relation that is sometimes called predication, a relation between the fact or trope and the universal whose existence the existence of the fact also implies. Some philosophers would say that the universal happiness is predicated of the fact or trope John being happy or John’s happiness. It is important to notice that these three kinds of ontological predication are entirely different, even though they may be definable with the aid of each other. While many philosophers use different expressions for the three kinds of ontological relations associates with linguistic predication, there is unfortunately no standardized terminology. Nevertheless the distinction has been made. E. g. when an entity is predicated of another in the first sense Brian Ellis says (see [Ell01, 1.3] that it is instantiated in it but when it is predicated of another entity in the third sense Ellis says that it is instantiated by it. Sometimes e. g. in Neo-Aristotelian ontologies such as Jonathan E. Lowe’s [Low06, page 18] ontological predication in the first sense is called instantiation and ontological predication in the second sense (a relation between particulars which Lowe calls objects and tropes which he calls modes) is called characterization and ontological predication in the third sense is called exemplification. Other philosophers, however, use these
expressions interchangeably or even in opposite senses (ontological terminology is very unsettled and confused!). I will use them interchangeably in this dissertation.

It is clear that the second ontological predication relation is not one such that the same entity could by it be predicated of many subjects. John’s being happy cannot be predicated of any other man in the second sense. If Mike is also happy, then the universal happiness is predicated in the first sense of Mike as well as of John, but the fact, John’s being happy, is not predicated of Mike in the second sense. Thus the second concept of predication does not give rise to any non-trivial concept of a universal. If it were the only non-linguistic relation associated with linguistic predication, then obviously there would not be universals. However, that this relation does not give rise to a concept of universal that could be satisfied does not of course exclude the possibility that there might be also another relation associated with linguistic predication, a relation between universals and entities that are connected with them; in the next section of this dissertation I will try to show that there is such a relation. In fact if there is such a relation there are many such relations. It would not conduce to clarity to call all of them by the same word, however, so I will as a matter of stipulation call only predication in the first sense, the relation that (if it exists at all, which is yet to be proved) holds between happiness and John when John is happy, and not between happiness and John being happy, nor between John’s happiness and John, instantiation.

There is one way in which the concept of instantiation used by realists about universals can be distinguished from the trope-theoretical concept of predication without assuming the truth of either trope theory or realism. According to the realist a word such as ”happy” denotes the universal happiness in all sentences where it occurs, in all endophoric contexts. However, the trope theorist cannot say that the word denotes any trope independently of context, but only dependently on a context. The trope theorist has to say that in the context of the sentence ”John is happy.” the word ”happy” denotes John’s happiness while in the context of the sentence ”Mary is happy.” it denotes a different trope, Mary’s happiness. Thus it seems that the concept of instantiation can be clarified by characterizing instantiation as the relation that is an objective correlate of linguistic predication and is such that the linguistic arguments and predicate denote the relata of this relation.
independently of context.

It might be thought that if they are to satisfy Aristotle’s definition, universals are those entities that are instantiated by many entities, and many philosophers have drawn this conclusion. However, it is important to see that this is not correct. Since as we have seen the Aristotelian conception of non-linguistic predication is the disjunction of instantiation and identity, even an entity that is instantiated by just one other entity (if there is any such entity, such as individual essences are claimed to be by many modern ontologists) is predicated of many entities in Aristotle’s sense, because it is predicated of two entities. It is not only predicated of the entity that instantiates it but also of itself, for since as it is identical with itself it is also predicated of itself in the Aristotelian sense of the word “predication”\(^\text{17}\).

Therefore once we distinguish instantiation/exemplification and identity, as a competent modern philosopher is bound to do, it is natural and almost

\(^{17}\text{This can be confirmed by considering one of Aristotle’s examples later in De Interpretatione. Aristotle says (21a):}

\begin{quote}
Of things predicated separately, some can be predicated in combination, others cannot . . . Further, if Socrates is a man and is Socrates he will be a man Socrates . . .
\end{quote}

Aristotle is here giving examples of things predicated both separately and in combination, and Socrates and man are one of his examples. Thus Aristotle says that man and Socrates are predicated both separately and in combination of Socrates, i.e. man is predicated of Socrates and Socrates is predicated of Socrates and the man Socrates is predicated of Socrates. It follows from this that Socrates is predicated of Socrates. Surely Socrates is not an exception, but Aristotle presupposes that every particular is predicated of itself. This might, indeed, seem to contradict what Aristotle says in Categories (2a10), where he says that primary substances (such as Socrates) are not said (\(\text{λέγεται}\\\text{, legetai) of any subject, as being said of and predication may seem to be the same. We cannot assume as medie\}}

\begin{quote}
\text{val philosophers did that Aristotle’s doctrines in De Interpretatione and in Categories are consistent with each other, so if there was a contradiction here this would not be fatal to my interpretation of what Aristotle says in De Interpretatione. However, there in fact need be no contradiction here between the two works. Aristotle apparently has two different concepts expressed by the two words, being predicated (\(\text{κατηγορηται}\\\text{, katetegoretai) and being said of (\(\text{λέγεται}\\\text{, legetai), even though they are often translated by the same word, “predicate” being often carelessly used also as an alternative translation of (\(\text{λέγεται}\\\text{, legetai), though this is usually translated by “say”. The concept of being said of a subject used in Categories is not the same as the concept of predication used in De Interpretatione (and also occasionally mentioned in Categories) but a subconcept of it. In the terminology of modern logic, \(\text{κατηγορηται}\\\text{ stands for a reflexive relation while \(\text{λέγεται}\\\text{ (at least in combination with the words καθ υποκειµένων, kath hypokeimenou, “of a subject”) stands for an irreflexive one, at least as Aristotle uses the words in the two works mentioned or at the very least as he uses them in the places in them referred to. Even if Aristotle does not always keep the two terms carefully apart, yet we can distinguish the two concepts. Aristotle’s definition of universals, that has been the basis of the historical discussion of universals, is in any case based on the wider concept Aristotle uses in De Interpretatione which takes identity to be a subspecies of predication.}

\end{quote}

430
inevitable to define the concept of a universal so that it applies to anything that is instantiated, whether by one entity or many entities, so long as the instantiation relation or predicate appealed to in the definition itself is of such a kind that it could connect an entity to many universals.\footnote{This conclusion can be independently confirmed by what Porphyry in \cite{Por92, page 68} says of species, which are of course by Aristotelians taken to be a kind of universals. Porphyry says that it does not hold in general that a species is predicated of several things differing in number, and gives as a counterexample the bird species phoenix, of which it was thought there was at a single time only one specimen. This example may seem funny to modern philosophers, that do not usually believe in phoenixes; however, the logical point remains valid, even if the example must be taken as fictional and not factual. However, this point is not as strong as the previous one; the species phoenix can according to the myth be instantiated by different particulars at different times, while an individual essence would always have to be instantiated by the same entity.}

Of course it is not necessary to use the word in this way; one could also restrict the word “universal” so that it applied only to those entities that are instantiated by many entities and coin some new word, such as Wolterstorff’s \cite{Wol70, page 65} ”predicable entity”, for what I call universals. However, the word ”predicable entity” could be misleading as it could be taken to refer to a linguistic expression. This could be remedied by modifying it to ”ontically predicable entity” or ”metaphysically predicable entity”. However, as it is in any case inconveniently long as a name for a concept I will need to use frequently, I choose rather to extend the meaning of the word ”universal”. Obviously neither choice of words is right or wrong but this is a matter of arbitrary decision; however, it is clearly both the most convenient decision and the one that accords best with the historical tradition once this is understood rightly.

Universals must not be defined as those entities instantiated by particulars, since this would arbitrarily exclude the possibility of universals instantiated by other universals, higher-order universals, but as those entities instantiated by some entities, not matter what entities.

Furthermore, if it is accidental for some properties which entities instantiate it or if it is instantiated at all, then it is natural to extend the denotation of the word ”universal” even further to apply to any entity that could be instantiated, since a term expressing an ontological category should apply to any entity to which applies necessarily. This definition leaves of course open whether there actually are uninstantiated universals.\footnote{Paul Vincent Spade translates \cite{Spa96, page 145} the passage ”I call universal that which is naturally apt to be predicated of several things”. This translation would seem to suggest that even according to Aristotle’s definition a universal need not be actually}
Instantiation is often taken to be an asymmetrical relation, and it has been thought that it must be asymmetrical for this definition of universals to define an interesting notion. I am not assuming that it is asymmetrical, however. This would make self-instantiation impossible, but it seems intuitively to be possible; e. g. it seems natural to say that the property of being a property is a property and is therefore instantiated by itself. Considerations relating to paradoxes might of course make us to retract this intuition, but there are solutions to the paradoxes that allow us to hold fast to it. In any case, it is enough for the definition to define an interesting concept that the relation of instantiation be non-symmetrical. I. e. it need not hold by definition for every entity that whatever it instantiated it is also instantiated by (though of course for all we know at this point in our investigation this might hold).

The concept of a particular has been usually understood as the opposite to that of a universal. Particulars can thus be defined simply as entities that are not universals, i. e. that could not be instantiated by any entity. Of course this definition also leaves open whether there are any particulars; for all we can say at this point in our investigation, it could be that all entities were universals (or if there were no relation of instantiation it could even be that there were no universals and no particulars either; both notions might fail to apply to anything).

Porphyry himself later in Isagoge makes a distinction between predications which are said of only one thing and others said of many things. Porphyry tells that the first group includes besides genera and species also differences, properties (here the word is used very differently from the way modern philosophers commonly use it) and common accidents (συμβεβηκότα κοινως). All these are traditionally called predicables20.

20The notion of a predicable comes originally from Aristotle’s Topics. Porphyry’s list of predicables, however, is different from Aristotle’s original one, since it includes species as predicables while Aristotle had definitions in their place. Many commentators treat Porphyry’s divergence from Aristotle as a major error, but I cannot see any reason for this. Etymologically both the notion of category and the notion of predicable come from the same source, as both notions signify things predicated in different ways of something, so they are easy to confuse, especially as “accident” is both a name for a predicable and a common name for all categories other than substances. Scholastic philosophers called categories predicaments to distinguish them from predicables.

432
Common accidents are predicated accidentally, while differences and properties apparently are predicated necessarily of whatever they are predicates and differ from genera and species in some other way (apparently in not carrying a principle of individuation, whatever this means). Obviously all the same questions Porphyry asked of species and genera can also be asked of these predicables, of differences, properties and common accidents, though Porphyry did not do so. We can ask whether common accidents are real or are situated in bare thoughts alone, whether as real they are bodies or incorporeals and whether they are separated or in sensibles and have their reality in connection with them. It seems best to include all questions like these in the problem of universals even in its traditional form, since surely the problem of universals should be a problem that concerns all universals, not just some subset of them.

Later discussion has in any case commonly concerned also entities predicated accidentally of something else. Indeed, Armstrong would in [Arm97, §3.9, page 44] only give the name of universals to such properties that when they are truly predicated of a particular the resulting truth is a contingent one; he calls other properties third-class properties. Thus the properties Armstrong would call universals are clearly properties predicated non-essentially of something. They are thus like Porphyry’s common accidents (or more exactly Porphyry’s separable common accidents), while Armstrong would call Porphyry’s species, genera, differences and properties alike third-class properties (if he allowed them to exist at all). Thus the entities the problem of universals would concern according to Armstrong and those it would concern according to mediaevals are mutually exclusive. We could then argue, as Markku Keinänen does in [Kei05, page 102] that there are two distinct problems of universals, which he calls the problem of kind universals and the problem of universals.

However, these two problems are nevertheless not wholly dissimilar; what is common to both is that they both concern entities which are metaphysically predicated of something. Many modern philosophers would follow
Porphyry in calling essential properties universals, so it does not seem good to follow Armstrong in restricting the discussion to non-essential i.e. accidental properties. To encompass both the medieval and modern discussion it seems to me best to formulate the problem of universals so that it concerns whether anything predicated of anything (in a non-linguistic sense) exists at all. If the answer to this question is negative then neither species and genera in Porphyry’s sense nor universals in Armstrong’s sense exist, so this is the most fundamental question; if the answer is positive then we can ask further whether essentially predicated or accidentally predicated entities or both exist.

As Spade shows in [Spa94, page x], universals were also defined as what is common (κοινον, koinon) to many things, where the way in which they are common was made precise at length in many ways. This use was introduced into medieval philosophy in Roman Neoplatonist Boethius’s second commentary on Aristotle’s De Interpretatione, where Boethius used an earlier commentary on the same work by the Greek Neoplatonist Porphyry. For example, Porphyry and Boethius specified that it did not mean being common in the way a horse or a slave could be common to two brothers or a public bath was common. A universal was according to Porphyry’s definition (see [Spa95, page 42] and [Por92, page 42]) common in the sense in which that is called common which, as a whole, comes undividedly into the use of many simultaneously. Because of this, S. Albert Kivinen has in [Kiv99] called the theory that there are universals ontic communism - to be clearly distinguished from political communism, of course!

Another useful way of using the concept of universals would of course be to make this notion of common entities precise by defining universals as those entities that are constituents of spatio-temporally separated entities. Many, perhaps most modern ontologists that use the notion of universals follow this course. While this notion is in my view quite coherent, it is done). Plato did not yet make a distinction between accidental and essential predication central to his philosophy; as both Loux and Lewis emphasize, this distinction or at least its centrality was an innovation of Aristotle’s (though later Platonists picked it up). Many modern commentators have held that Plato saw all Ideas as predicates accidentally of (participated accidentally in by) sensible particulars; probably the only essential predication in Plato’s theory was the idea’s self-predication (if Plato indeed held ideas to be predicated of themselves, which has also been denied) or predication whose subject is a soul. Many of Plato’s examples of Ideas appear to be accidental properties or entities possessing accidental properties.

434
less general than the notion I have defined (as Wolterstorff shows clearly in [Wol70, chapter 10]). Therefore it is less suited to be a basic concept of ontology. Besides if we used it we would exclude some theories that have generally been viewed as answers to the problem of universals from being relevant to the problem; we would have to say that transcendent realism (such as has often been ascribed to Plato and the Neoplatonists) would not be a form of realism concerning universals, as they take universals to be entities outside spatio-temporal particulars, which seems unnatural\(^\text{23}\). It seems to me better to use a snappy and well-known term like "universal" rather than a clumsy expression like "ontically predicable entity" for a basic ontological concept. While the way I use the term "universal" may be a minority use, yet as we have seen it has at least as long a history behind it as the more common use, and other important philosophers of today such as Wolterstorff, Loux and Bealer have also used the term similarly to the way I use it, so I am not just arbitrarily deviating from established linguistic conventions. There is no single right way to use a technical term; one must just choose for reasons of convenience which of the current usages to adopt.

There is a third way in which universals are commonly defined, namely be reference to their relations to space-time. This can be done in two very different ways. Immanent realists often take universals to be entities that could be simultaneously present in a single spatial location, while particulars would be entities that could only be present in a single place at once. This definition of universals may take its origin from the second medieval idea of a common entity. Transcendent realists, however, often define universals as entities that are not located in space and time at all and particulars as entities that are located in space or in time. This distinction is often marked also by use of the terms abstract entities and concrete entities (though the terms "abstract" and "concrete" are also often used by trope theorists in a very different sense to stand for dependent and independent entities).

\(^{23}\) Gustav Bermann claimed that Husserl was a nominalist on the basis of this definition, and such an ontologist as J. P. Moreland in [Mor89] has seriously discussed on the basis of Bergmann’s claim whether Husserl was a nominalist. This results from accepting this narrow definition of universals. However, it seems clear to me that Husserl was a realist about universals if anyone, whether or not he was a transcendent or immanent realist (which is indeed a difficult question).
In most cases the immanent realist variant of this definition leads to the same results as the first definition of universals. However, as many philosophers such as both Wolterstorff and Rosenkranz notice, there appear to be possible examples of entities which would be universals by this definition but not by the first one. It has often been thought that there could be disembodied spirits, which were nowhere. Such spirits would be neither universals nor particulars by this definition. However, including spirits and universals in one category would make that category rather heterogeneous. Surely it is more natural to classify such spirits as particulars.

6.3.3 Different Kinds of Purported Universals

Abundant and Sparse Universals

Rodriguez-Pereyra also argues that the problem of universals is about the properties which David Lewis in [Lew86, page 59] calls sparse or natural properties, not about all abundant properties. Lewis says:

Sometimes we conceive of properties as abundant, sometimes as sparse. The abundant properties may be as extrinsically, as gruesomely gerrymandered, asMiscellaneously disjunctive, as you please. They pay no heed to the qualitative joints, but carve things up every which way . . . The sparse properties are another story. Sharing of them makes for qualitative similarity, they carve at the joints, they are intrinsic, they are highly specific, the sets of their instances are ipso facto not entirely miscellaneous, there are only just enough of them to characterize things completely and without redundancy.

While Porphyry and the medievals of course did not possess Lewis’s distinction between sparse and abundant properties, it is likely that Porphyry and the medievals did intend the problem to concern only such properties as Lewis would classify as sparse. However, it seems to me that abundant properties also satisfy Aristotle’s definition of a universal (as it is most naturally understood), as they too are predicated of many things. We can surely ask

\footnote{Compare what Gail Fine says in [Fin80, page 210] of koina, i.e. common entities: koina are what Aristotle elsewhere calls "universals" (katholou).}
many of the same questions about abundant properties as we can ask about sparse ones; we can ask whether in the first place there are any such and if there are, are they independent of the human mind and language; Lewis himself thought that the answer to both questions was positive but many philosophers would disagree. Also most philosophers who identify themselves as nominalists would not be satisfied with the statement that there are no sparse properties but there are abundant properties; they would think that this claim is incompatible with their nominalism and amounts to a form of realism, and thus understand realism as a weaker claim that Rodriguez-Pereyra does. Thirdly the question whether or not abundant properties exist is in any case a genuine ontological question, which ought to be addressed in ontology, whether in connection with the problem of universals or separately. However, many though not all of the arguments for or against the existence of non-natural properties are very similar to, perhaps even of the same form as, those for or against the existence of natural ones, so it is useful to discuss them together and for this cause join them into a single problem. Indeed, I will argue that there are reasons to think that arguments

and what he says in [Fin80, page 210]:

Aristotle believes that the OMA shows that there are koina. He also believes that it allows that negations may be predicated. Hence he appears to be committed to there being koina of negations.

Thus if Aristotle is committed to there being koina of negations, and if koina are the same as universals, then Aristotle is committed to universals of negations, i. e. negative universals. As such negative universals would in modern philosophical terminology be abundant, not sparse, properties, it follows that Aristotle should in the interest of consistency call such non-sparse properties universals, though he might be unwilling to do so. It must be noted that though Fine points out this puzzle in Aristotle’s view, she would probably not herself be willing to accept the conclusion that there are negative universals, as she follows Armstrong in favouring a sparse realism. However, she does not point out any way to avoid the conclusion, and therefore she unwillingly provides evidence for promiscuous realism.

If the question regarding the existence of abundant properties would not be a solution to the problem of universals, and the assumption that there are abundant properties and they are independent of the human mind and language would not amount to a form of realism, we could introduce a new form of nominalism alongside the classical ones distinguished by Armstrong: abundant property nominalism! This form of nominalism would reduce apparent talk about sparse properties (which would be universals) into talk about abundant properties (which would not be universals) and so get rid of apparent universals. Indeed, if I came to believe that it was correct to use terminology in this way, then I would consider abundant property nominalism a highly plausible form of nominalism, and would likely become an adherent of this theory and renounce my allegiance to realism. However, if the idea of abundant property nominalism sounds funny or even absurd to you, this is surely a reason to consider the theory that there are abundant properties as a form of realism concerning universals.
for the existence of sparse properties must use a premise stating that there are abundant properties and so go through abundant properties. Fourthly there are some serious (even if not obviously insuperable) problems with the notion of a sparse property.

Lewis suggests many different criteria in the above quotation. It is problematic how they are supposed to be related. They can surely not be assumed to be equivalent unless some very strong assumptions are taken for granted. Nor is it even obvious that putting all of them together yields a coherent concept of a property.

Also some of the criteria are slightly obscure. The expression "carve at the joints" is obviously metaphorical so it is not very helpful. Besides, it raises a question a radical promiscuous realist may well pose (in equally metaphorical terms) - what if there are no joints in nature? In that case an objective division of properties into sparse and abundant would be impossible.

The more literal criteria also pull in different directions. Especially, I see no reason why the properties sharing of which makes for qualitative similarity would have to be highly specific and such that there are just enough of them to characterize things completely and without redundancy. It seems to me clear that if two properties make for qualitative similarity and are compatible then their conjunction also makes for qualitative similarity; however, this already means introducing redundancy, for whatever can be characterized with the aid of the conjunction of the two properties can also be characterized with the aid of the two properties themselves.

More worryingly the criterion that the sparse properties are maximally specific is unclear and fits badly with the rest of Lewis’s non-metaphorical criteria. It is not clear how specific sparse properties would have to be; if it is taken literally that they must be maximally specific, then every individual has just one sparse property, the conjunction of all its intrinsic and qualitative properties. These kinds of conjunctive properties are sparse indeed, but this is surely not what Lewis intended. They are not likely to satisfy the criterion that there are just enough of them to characterize things completely and without redundancy. It must also be noted that this conception of properties would surely be no more natural than the conception of abundant properties. We surely do not ever quantify over that kind of properties in everyday speech! Rather, it seems Lewis merely means that no sparse
property is a subproperty of another sparse property. However, if maximal specificity is understood in this way, then many different concepts of properties that require the properties that satisfy them to be intrinsic and qualitative may satisfy the criterion.

Lewis probably does not intend to give a reductive definition of sparsity. Rather, the criteria are only intended to give a guide to readers’ intuitions, helping them to pick out a primitive, intuitively most obvious concept of a sparse property. However, it is by no means obvious that there is only a single intuitively most natural concept of properties satisfying the criteria. We must take into account the possibility that philosophers who say they are discussing sparse properties vacillate between different non-equivalent conceptions of properties without realizing it. Looking at how other philosophers than Lewis characterize sparse properties will make this possibility a near certainty.

In my view the best reason for supposing that there are important sparse kinds of properties is found in Goodman’s new riddle of induction (see [Goo54]). Goodman saw that predicates had to be divided into those that are projectible and those that are not. A predicate "P" is projectible (with respect to a predicate "Q") if from the fact that it has been observed to be applicable to all individuals that are Q we have observed we can inductively infer that it is applicable to all individuals that are Q. Goodman gave as a famous example of a predicate that is not projectible "grue", defining this so that it applies to all things examined before t just in case they are green but to other things just in case they are blue (here t is some arbitrarily chosen time). Though as a nominalist Goodman himself did not do so, many philosophers have since generalised the distinction into a distinction between properties besides predicates, a distinction between those properties that are projectible and those that are not. A property P is projectible (with respect to a property Q) if from the fact that all individuals of kind Q have been observed to be of kind P we can inductively infer that all individuals of kind Q are also (probably) of kind P.

Carnap tried to solve Goodman’s riddle in [Car71, pages 72-76] by supposing that primitive projectible attributes (i.e. as Carnap uses the word "attribute" properties and relations) are qualitative attributes. In Carnap’s terminology they must be descriptional (as contrasted with locational); Carnap himself used the word "qualitative" in a narrower sense than is usual in
the present literature so that it stood only for properties that were neither spatial nor temporal (e.g. shapes and durations would not be qualitative in Carnap’s sense, though they would often be called qualitative). Qualitative properties are usually understood as properties that make no reference to any specific particular. Carnap defines his locational attributes so that they serve to specify the absolute, not only the relative, location of an object or event in space and time, and that they serve to identify an object or event. A primitive attribute is descriptional iff it is not locational. A possible obscurity in Carnap’s theory is what it means that an attribute is primitive and what reason Carnap has to think there are primitive attributes. In any case according to Carnap primitive attributes must be chosen so that they can be grouped in families. A non-primitive attribute is presumably projectible if it can be defined solely in terms of descriptive attributes. The two characterizations coincide if we have a relational theory of space and time; if we have a substantivalist theory of space and time Carnap’s characterization is more general.

It seems to me that Carnap was clearly right in claiming that all projectible properties are qualitative. Obviously grueness is not qualitative; it is formed from the primitive attribute of being examined before t which is locational in Carnap’s sense. However, some philosophers have presented examples that seem to show that the converse need not hold, that is, that some qualitative properties need not be projectible. I will leave open the question of whether this is so or not.

It must be noted that not all projectible properties need fulfil Lewis’s criteria for sparse properties, since at least some relational properties seem to be projectible, and thus all projectible properties need not be intrinsic. For instance, if all mouses have been observed to be smaller than elephants, we can inductively infer that all mouses are smaller than elephants. Therefore being smaller than an elephant seems to be a projectible property. Beings smaller than an elephant, however, is apparently a relational, not an intrinsic property, though it is a qualitative/descriptive relational property, since it involves reference only to elephants as a kind, not to any specific elephant.

However, though the distinction between projectible properties and other

---

Both concepts of qualitative properties are of course narrower than the sense of the word "qualitative" in which qualitative properties are contrasted with structural ones. Neither qualitative properties in Carnap’s sense nor locational properties are structural, i.e. invariant under all isomorphisms.
properties cannot be conventional, it is not clear that it has ontological significance. The distinction is obviously primarily an epistemological distinction, as it is formulated in epistemic terms; a projectible property is distinguished from other properties by the way in which it can be used in inductive inference. While an ontologically important distinction might of course underlie and explain this epistemological distinction, no one has so far as I can see given any good reasons that this would be so. It may well be so that properties are not projectible or not-projectible in themselves, but only relative to the cognitive faculties of a person. We cannot decide arbitrarily which properties to project, as one choice leads to justified and another to unjustified beliefs, but the reason for this may be as much in our own nature as in the nature of properties themselves. It may be that different properties would be projectible to different kinds of intelligent creatures. It might be for instance that a property is projectible only if its instantiation can be directly observed in some cases. We can perceive immediately that something is green without perceiving it to be grue, and perceive immediately that something is blue without perceiving it to be grue but we certainly cannot immediately perceive it to be grue without perceiving it to be green or perceiving it to be blue. This might explain the epistemic asymmetry between greenness and grueness. If this is the right explanation then if different kinds of creatures can perceive different properties immediately then different properties would be projectible for them. In this case the distinction between projectible properties and other properties would not have ontological significance.

Another criterion for distinguishing sparse properties is that sparse properties are those that are mentioned in laws of nature, or are denoted (or are some kind of semantic value of) basic terms of basic scientific theories. Let us call this kind of properties nomological properties. It is usually supposed that projectible properties are the the same as those that occur in laws of nature. However, I do not see any good reasons for accepting this supposition, and in fact there are reasons for being very suspicious of it. We can surely make low-level inductive generalizations that involve properties that need not be mentioned in laws of nature. All general conclusions arrived at inductively need not be laws of nature, but there can also be accidental generalizations whose truth depends on the initial conditions of the universe and not only on laws.
Indeed, I cannot see why occurring in a natural law would be a criterion that would pick out any very interesting conception of property. There is a reason to be very suspicious of the whole notion of a nomological property. It is generally accepted that the same scientific theory can have different equivalent formulations and that the same natural law can be expressed in many different but equivalent ways. Because of this it is usually held in the philosophy of science that scientific theories (all scientific theories, and thus also basic ones) are deductively closed in the sense that all logical consequences of the statement contained in a theory are also contained in the theory. There are indeed reasons to be suspicious of full deductive closure (under standard logic) but it is scarcely dubitable that some kind of deductive closure, even if limited, is essential to a scientific theory. Thus different systems of axioms can define the very same theory. It seems to me that if this is so then since basic concepts or terms are analogous to axioms, scientific theories would also have to be definitionally closed; that is, every concept definable in terms of those occurring in a formulation of a theory would also occur in that theory. Two formulations of a theory with different basic concepts or basic terms would define the very same theory. However, if this so then occurring in a basic scientific theory cannot be a criterion that could distinguish sparse properties from abundant ones in any way similar to the way ontologists usually distinguish them.

If scientific theories are definitionally closed then the disjunction of two nomological properties is also a nomological property; thus disjunctive properties would be sparse, though nearly every philosopher who draws the distinction thinks them to be abundant, not sparse. Likewise the conjunction of two nomological properties would also be a nomological property and the negation of a nomological property would also be a nomological property. Likewise the property of being self-identical can be defined with the aid of identity, and surely every non-trivial theory must contain the concept of identity, and thus if theories are definitionally closed the property of being self-identical must be mentioned in every theory. However, self-identity would be held by every philosopher who uses the term to be an abundant and not a sparse property. On the other hand if one rejects the definitional closure of theories then it becomes a problem how different formulations of the same theory are to be identified. If we reject definitional closure we should reject also deductive closure. This would require a completely new
theory of individuation for scientific theories and in essence a quite new philosophy of science. It may not be impossible to construct such a theory, but I think we should be sceptical until someone succeeds in doing so.

Thus nomological properties are not very good candidates for sparse properties. Projectible properties are better ones, since they at least cannot be definitionally closed; however, they may be subjective in a way that nomological properties are not and this may rob them of their ontological significance.

Armstrong himself makes distinctions which may be analogous to that between abundant and sparse properties in his later work when he distinguishes in [Arm97, page 44] between first class, second class and third class properties and relations. According to him second class properties are distinguished from third class properties by the fact that when truly predicated of a particular, the resultant truth is a contingent one.

However, Armstrong does not make wholly clear what distinguishes the first class properties which he calls universals from the second class ones. Armstrong says that second-class properties require second-class states of affairs and that second-class states of affairs supervene on first-class states of affairs, which have first-class properties as their constituents. Armstrong also argues that the second class and third class properties and relations are no addition of being to the first class properties; nevertheless they exist and must be formally distinguished from the first class properties and relations. Apparently Armstrong then thinks that first-class properties are those properties that are constituents of fundamental states of affairs where the sense of fundamentalness in question is one that can be defined with the aid of the concept of supervenience. This would imply that first-class states of affairs would not supervene on anything that does not supervene on them but second-class states of affairs do; Armstrong does not explicitly say this, but I do not see how Armstrong’s characterization of first-class states of affairs could pick out a unique class of states of affairs nor thus how first-class properties could pick out a unique class of properties unless we assume that this is what he means.

However, in this case if fundamentalism is false, then all states of affairs

\textsuperscript{27} Schaffer has already pointed out in [Sch04, page 96] that there is a tension in Armstrong’s conception of universals i.e. first-class properties and relations as he sometimes seems to think of them as only fundamental properties and sometimes seems to use a broader conception.
may be second class states of affairs and hence all properties may be second class properties! It may be that for every state of affairs there is a state of affairs on which it supervenes but which does not supervene on it. Armstrong does not take this possibility into account; he considers the possibility that every property might supervene on other properties, but not the possibility that this might also happen for states of affairs.

Armstrong restricts the term "universals" to his first class properties. It is of course a matter of stipulation how such technical terms as "universals" are to be understood; however, this terminological choice is by no means the only natural one, and can even be viewed as misleading. We saw that universals are historically often understood as those entities that are predicated of many entities and that it is natural (once we distinguish instantiation and identity, as a modern philosopher is bound to do) to extend the meaning of the word so that it applies to anything that could be instantiated, whether by one entity or many entities. Many of the abundant properties that are not also sparse properties, the second and third class properties, are of course instantiated as well as sparse ones, and many of the abundant properties are instantiated by many entities as well. For instance, since many individuals are identical with themselves, many individuals instantiate the property of being self-identical. As Spade shows, universals were also defined as what is common to many things, where the way in which they are common was made precise at length in ways I cannot here go into. This does not lead us as readily to think of universals as abundant, but I do not think even this this definition of universals is incompatible with regarding them as abundant properties and relations. It is of course true that in the Greek and scholastic discussion of universals the attention was nearly always focused on universals that were in some very strong though never clearly specified sense natural, though this did not follow from the definitions Aristotle and many of the scholastics gave, so the terminological choice is not determined uniquely even on purely historical grounds.

However, I will in this dissertation stipulate that I use the word in the broader sense as a common name for abundant properties and abundant relations. This broader sense seems more useful to me, as it is likely that there are no first-class properties and hence no universals in Armstrong’s sense, but in that case there may still be universals in the sense I use the word. In that likely case the word "universal" as used by Armstrong will
not denote anything and will be thus rather useless, but as used by me it will even in that case denote something.

This distinction between abundant and sparse properties seems very useful to me, at least as a preliminary distinction, though adopting it does of course not imply that we should adopt Lewis’s highly controversial theory of what abundant properties are (nor of course his theory of what sparse properties are either). Lewis thinks that abundant and sparse properties are exclusive categories, though some abundant properties, natural ones, can correspond to sparse ones. It seems simpler to me to take sparse properties to be a subcategory of abundant ones and identify sparse properties with natural properties. Lewis thinks that abundant properties and relations must be analysed as sets of possible individuals, while he is willing to accept (in [Lew86, page 63]) sparse properties and relations as primitives. Indeed, I suspect Lewis has to take sparse relations to be primitive. Lewis uses mereological and spatio-temporal relations to define worlds. Lewis, however, tries to reduce abundant relations to classes of their possible instances. However, trying to apply this method to mereological and spatio-temporal relations would apparently be circular. Therefore Lewis must take at least these relations as primitive. Lewis also has to take sets as a primitive in order for his reduction of abundant universals to get off the ground.

Lewis, however, pejoratively calls theories that take abundant properties as primitive and try to analyse modalities with their aid “magical ersatzism”. However, Lewis provides no reason why one conception of properties could be taken as a primitive while the second would have to be analysed! Lewis writes [Lew86, page 189]:

‘Property’, and the rest, are names associated in the first instance with roles in our thought. It is a firm commitment of common sense that there are some entities or other that play the roles and deserve the names, but our practical mastery of uses of the names does not prove that we have much notion what manner of entities those are.

Cannot the same be said of sets and spatio-temporal and mereological relations? Are not they too associated in the first instance with roles in our thought? Why would we have any more notion of what they are than we have of what properties are? Cannot a theorist that takes properties as prim-
itive complain that Lewis does not explain what sets and spatio-temporal and mereological relations are? Indeed, van Inwagen has already posed just this objection to Lewis’s argument. Are not spatio-temporal relations and mereological relations and sets taken as primitive just as much (and just as little) magical as abundant properties taken as primitives? Indeed, sets seem to be in a worse position than properties in one respect; it is not a firm commitment of common sense that there are any entities that play the role of sets and deserve the names. That there are sets may be a commitment of highly developed mathematics and science but not of common sense. That science is committed to them is of course a reason to believe in them, but arguably not as strong as the reason we have in believing in properties. To be sure, the properties that common sense is committed to are primarily those close to sparse properties; however, they are surely not all quite sparse in Lewis’s sense; e.g. people have always quantified over such properties as families and professions, and these are surely not intrinsic, as sparse properties would have to be, but clearly relational; e.g. two people belong to the same family if they have the same ancestor or one of them is an ancestor (or mother or father) of the other.

Later in [Lew91a] Lewis admitted that sets are a problem for him. He proposed that the subset relation could be understood as the mereological part-whole relation; however, this leaves the relation between a singleton set and its sole member unanalysed, and Lewis admits that he comes to an impasse here. He finds the notion of a singleton profoundly mysterious, yet cannot bring himself to abandon standard set theory because of this. He toys with the idea of eliminating sets. He draws on results by Burgess and Hazen according to which (see [Lew91a, pages viii, 121] we can get rid of sets supposing that there are infinitely many atoms and not too much atomless gunk. These are of course quite unjustified suppositions. In the end Lewis himself did not find this satisfactory either. Yet he never considered taking the notions of instantiation and abundant properties as primitive as a way out.

It may be argued that it is surely better to have less than more unanalysed entities. This is true. However, we need not have more unanalysed entities even if we hold abundant properties to be primitive entities, for it can also be suggested that we can reduce sets to abundant properties.

In any case for all these reasons it seems to me that it is useful to
formulate the problem of universals more generally so that it concerns both kinds of properties.\footnote{There is a further reason for the more general formulation. Rodriguez-Pereyra just states in [RP02, §1.1, page 20] that the resemblance which accompanies sparse properties is ontological and objective. However, other philosophers may deny this objectivity - as Rodriguez-Pereyra admits that Goodman denies it - and if so may question whether the distinction between sparse and abundant properties is at all objective. Such a philosopher would then have to think that the problem of universals as Rodriguez-Pereyra formulates it is not a legitimate problem, but many ontologists would surely think that such a philosopher would offer a solution to the problem of universals rather than rejecting it. Surely Goodman is usually considered to offer a solution - namely a nominalistic one - to the problem rather than rejecting it as illegitimate. A philosopher may even reject the objectivity of this distinction without rejecting the objectivity of the distinction between particulars and properties of both kinds. Such a view would be a form of promiscuous realism concerning universals, and I myself am attracted to it. It is not good for the formulation of a philosophical problem to have such strong and controversial presuppositions as the presupposition that the distinction between sparse and abundant properties would be objective. If a philosopher thinks that the distinction is objective he must try to argue for this claim.}

**Are Sets and Classes Universals?**

It is controversial whether the problem of universals also concerns sets and classes. The terms ”realism” and ”nominalism” are used in very different and sometimes even quite opposite senses. A striking example of the prevailing terminological chaos can be seen in comparing the theories and terminology of Nelson Goodman and Armstrong. Goodman in his early theory in the *Structure of Appearance* defined the terms so that apparently according to him a theory is nominalistic if according to it there are no classes, even if it allows that there are properties. Goodman’s own theory which he called nominalistic is like this, since according to it [Goo51, page 156] though there were no classes there were qualia, properties of experiences (only experiences, since Goodman was a phenomenalist though a tentative one) like colors and sounds (though shapes and sizes were not qualia), that were actually components of several particulars, of several experiences. Goodman explicitly says [Goo51, page 158] that

the whole quale appears in every concretum in which it appears at all

Goodman does say that a nominalistic theory does not allow classes or other non-individuals, but it is not clear what entities other than classes would
be non-individuals according to Goodman. Goodman distinguishes indi-
viduals and particulars, but what this distinction amounts to is not clear
to me, nor has it been clear to most other philosophers, other than that
classes are not individuals and properties are not particulars according to
him. Of course later Goodman has become suspicious even of properties
while he drifted from phenomenalism to a totally subjectivist theory where
men are supposed to be world-makers. For the early Goodman (see [Goo51,
page 107]) realism was not the opposite of nominalism but of particularism,
but Goodman’s particularism is what Armstrong would call nominalism.
Thus Goodman’s nominalistic theory was also what Armstrong would call
realistic, while Armstrong’s realistic theory of properties and relations is
compatible with what Goodman called nominalism (though Goodman could
not have accepted Armstrong’s states of affairs for the same reason because
of which he thought could not accept classes, namely that their existence is
incompatible with extensional mereology.). On the other hand e. g. the the-
ory the late Quine sometimes supported in his more realistic moods, which
allows classes but not intensional entities like properties, would according to
Armstrong be nominalistic but not according to Goodman, who would call
it Platonistic. As I use the term Quine’s theory, Goodman’s early theory
and Armstrong’s theory are all realistic in the minimal sense, though all are
in my view too close to nominalism as all allow too few different kinds of
universals to really work.

If we look at how the founders of mathematical logic originally infor-
mally explained the meaning of the membership sign, we see that they

29 The word “individual” is very ambiguous. Different philosophers use it very differently.
The great difficulty with this word is that many of the senses given to it define a coherent
and interesting concept only given some substantive and controversial assumptions. One
of the most common senses is that given in *Principia Mathematica* [WR63, page xix],
according to which

An "individual" is anything that can be the subject of an atomic proposition.

Unfortunately, this notion defines an interesting concept only in a type-theoretical frame-
work like that of *Principia Mathematica*. In a type-free higher-order logic anything, any
entity, can be the subject of an atomic proposition, so this notion is completely trivial in
a theory based on such a logic. The sense of “individual” that seems most useful to me is
that in which it is synonymous with “particular” in the most common and useful sense of
that word and is contrasted with the notion of a universal. A particular or individual in
this sense is anything that can only occur in the subject position of an atomic proposition.

30 It is well known that the axioms of set theory cannot by themselves determine the
meaning of set-theoretic expressions. Opponents of set-theoretic realism have often (e. g
in [Fields]) used this fact to argue that set-theory does not have a determinate meaning.
However, we can instead argue from this fact that such informal explanations as the ones
connected it with predication. Whitehead and Russell explicated the membership predicate $\in$ in *Principia Mathematica* [WR63, page 25] as follows:

"$x \in \alpha$" may be read as "$x$ is an $\alpha$". Thus "$x \in \text{man}$" will mean "$x$ is a man" and so on.

In this they were following Giuseppe Peano, who said\(^{31}\) in [Pea89, page x]:

Signum $\in$ significat est. Ita $a \in b$ legitur $a$ est quoddam $b$ . . .

Both Peano and Russell and Whitehead then thought\(^{32}\) that set-theoretic membership was a kind of predication, and it follows from this that sets (with the possible exception of the empty set, which can probably anyway be treated as fictional) according to them satisfy Aristole’s definition of universals.

There is an additional reason to take membership to be predication, which does not depend only on informal explanations. It is commonly assumed in extensional semantics (as we have seen in discussing Tarski’s theory) that an atomic sentence is true if the sequence of the interpretations of its terms belongs to the set that is the extension of its predicate (i. e. in symbols using the kind of notation I have been using in other parts of this work, $P(t)$ is true iff $\|t\|_g \in \|P\|_g$, where $\|t\|_g$ is the interpretation of the term $t$ relative to the assignment $g$ and $\|P\|_g$ is the interpretation of the predicate $P$ relative to $g$). However, making this assumption amounts to treating set-theoretical membership as metaphysical predication, if set-

---

I cite below play an essential role in determining the meaning of such expressions, making the meaning of set-theoretic expressions at least more determinate than the axioms alone could do, whether or not it can become fully determinate. I have argued that we can see that some observable properties are predicated of some observable particulars, so we understand what predication is, and telling that membership is predication helps us to understand a bit better what it is and so what sets are.

\(^{31}\)This is translated in [vH77, page 89] as follows:

The sign $\in$ means is. Thus $a \in b$ is read $a$ is a $b$ . . .

\(^{32}\)It must be admitted that all founders of mathematical logic and set theory did not explain membership in this way; e. g. Cantor, Dedekind and Zermelo did not do so. This does weaken the force of my argument, but not remove it entirely. I also prompts concerns whether they were operating with the same notion of class at all as Peano, Russell and Whitehead. However, it would seem rather implausible that their concept of classes was a different one. At least, it cannot be said that Whitehead and Russell would have intentionally distorted the concept of membership because of their logicism, well as their explanation fits with logicism, since they were just following Peano in their explanation of what membership was, and Peano was not a logicist.

449
theoretical membership is taken as a non-linguistic notion, as a realist concerning sets would do. If this practice of semanticists is taken at face value, then we must take the membership relation as metaphysical predication and classes therefore count as universals in the traditional Aristotelian sense.

These appear to me to be good reasons to take classes to be universals. Therefore an ontological view which is realistic about classes also counts as realism concerning universals aka universalism. Therefore the ontological view Armstrong (in [Arm78, pages 28-43]) and others call class nominalism is in my view better described as a form of realism, set realism or class realism. Nevertheless, there are reasons to think it is only a very inadequate form of realism. The kind of universals to which common sense seems to commit us, such as properties and relations in intension, do not satisfy any axiom of extensionality as classes do, so common sense seems to be committed to other universals than classes, or at least classes composed only of actual entities.

6.4 The Problem of Universals systematically considered

We have so far seen only that understanding the problem of universals as a request for an explanation does not lead to a good description of how the problem has historically been treated. Showing that the modern view of the problem is not historically correct does not, however, decide whether it is correct as a normative prescription. It would of course be possible for those who think that the problem of universals is a problem about explanation to give up the claim that they are dealing with a centuries old problem and confess frankly that they use the words ”problem of universals” to denote a

\[33\] Armstrong himself considered whether the theory he called class nominalism would rather be a realistic theory, as it was an objectivistic theory. However, the reason I have for calling it a form of realism is not just that it is objectivistic, but that at least one common concept of classes satisfies the original Aristotelian definition of a universal.

\[34\] For instance, a shirt can change its colour even though other shirts remain the same colour, but if a shirt belongs to a set of shirts it continues to belong to it so long as it continues to exist and all the other shirts that are members of the set continue to exist. It might be possible to deal with such an example with the aid of sets alone if one subscribes to a B-theory of time, but such a dodge is not possible in the case of modal and counterfactual statements. A shirt that does not change its colour could change it and continue to exist even though other shirts did not change their colour, but a shirt could not fail to belong to the set of actual entities it belongs to if the other shirts continue to exist.
problem that Armstrong discovered a little less than thirty years ago. We must now ask if this would be a wise choice to make.

Of course, if there were something wrong with the problem as originally formulated - for instance if it was based on a false or questionable presupposition like the famous question "Have you stopped beating your wife?" - then it would be not only possible but necessary to replace it with a different, genuine problem. However, those philosophers who formulate the problem as a problem of explaining something do not typically attempt to show that there is anything wrong with the original problem, and indeed their theories typically imply an answer to the original problem. In fact Porphyry did a surprisingly good job in formulating the problem so that he minimized the presuppositions he made\textsuperscript{35}.

However, this does not yet settle the normative question. Even if the traditional problem of Porphyry is a genuine problem, it could still be that the modern problem is in some way a better subject for investigation. I must now ask if this is so or not.

I will now argue that understanding the problem of universals as a request for an explanation does not lead to any good prescription of how the problem should be treated either. I will do this in two stages. I will first rebut arguments given by Rodriguez-Pereyra and Sowyer that purport to show that the problem must be an abductive one. I will then show that the problem formulated by Armstrong is narrower than the traditional problem, and leaves out vitally important questions that must be addressed in an adequate treatment of the philosophical questions relating to universals. After this I will go further and suggest that there are reasons to think that the problem introduced by Armstrong may actually be a pseudo-problem, while there are no such reasons to think that Porphyry’s original problem would

\textsuperscript{35}There may be some unjustified presuppositions in Porphyry’s formulation, and if so we must get rid of them. Porphyry seem to assume that either all universals are corporeal or none are, and similarly that either all of them exist apart from sense objects or all of them exist in sense objects. It might of course be that some universals were corporeal and some not, and that some existed apart from sense objects while others did not. Similarly Porphyry supposes that if universals exist in sense objects then they are dependent on them. This may also, depending on how the words “in” (“$\epsilon n$”), and "dependent" are understood, be an unjustified presupposition. If the word “in” is taken to indicate constituency, then Porphyry presupposes that if universals are constituents of particulars, they must also be dependent on them. There is no reason to believe this, since many constituents of complex entities are not dependent on the wholes whose constituents they are.
be a pseudo-problem. These reasons may not be conclusive, but even if Armstrong’s problem were a genuine problem, the first part of my argument suffices to show that it is not fitted to replace the traditional problem.

No doubt many philosophical problems indeed have the form Nozick indicates. However, Nozick does not claim that all philosophical problems would have to be of this form, and we have seen that the problem of universals as it has been traditionally understood is not of this form. If some philosophers think that there are certain entities and other philosophers think there are no such entities\footnote{There are of course innumerable examples of such problems; indeed, there are no entities that all philosophers would agree do exist. Phenomenalists think that there are no physical objects but only sensations or sense-data, while physicalists think there are only physical objects but no sensations or sense-data and dualists think there are both physical objects and sensations; more modern anti-realists think there are no unobservable objects like electrons while scientific realists think there are unobservable objects like electrons and extreme scientific realists think that there are only such objects; endurantists think there are three-dimensional objects but no four-dimensional ones while perdurantists think there are four-dimensional objects but no three-dimensional ones; theists think there is a God while atheists think there is not; and so on and so on.}, this also already creates a philosophical problem: which of these two groups of philosophers are right? Are there such entities? This ontological question is of course in practice inseparably connected to the epistemological question: are there any good reasons to think there are such entities? Of course if it turns out that there really are no good reasons to think there are entities of those kinds, then there arises a new problem: why did it falsely seem to some philosophers as if there were such entities? Likewise if it turns out that there are those entities, then we must try to answer how can it be that most philosophers do not see it, especially if the evidence for their existences is rather obvious, as many realists would claim it is. However, before embarking on investigating this kind of problem it must be shown that there indeed is no good reason to think that there are those entities, else one will be begging the most important questions of all! Also, neither of these is likely to be a very serious problem; history shows us that human beings are incredibly adept at missing the obvious, as well as at seeing things that are not there!

Chris Swoyer suggests in [Swo96, page 248] that many philosophers who view their ontological arguments as demonstrative do so because they are in the grip of the demonstrative ideal, according to which

philosophy should proceed from utterly secure premises, by deductively valid steps, to utterly secure conclusions.
I agree with Swoyer that this would be an untenable ideal. However, one who thinks that an ontological argument is demonstrative need not believe in this kind of ideal. A demonstrative argument could only lead to a conclusion that would be utterly secure if we could be certain that its premises are true and that the rules of inference used are valid. However, neither is the case when it comes to ontological arguments or indeed any deductive arguments whatever. Both the premises of an ontological argument and the rules of inference used in it are typically controversial. Swoyer says that we can rarely find demonstrative arguments for or against philosophical views. This is not at all correct. I would say that we quite commonly find such arguments; however, this does not help to bring agreement or certainty, since the opponents of a view usually then deny the truth of the premises of the argument or sometimes, though more rarely, the validity of the logical rule used in the argument. Often even the person proposing the argument can be brought to doubt the truth of the premises or the validity of the logical rule. Even if the premises seem to be just matters of common sense, one can always argue that they are expressed misleadingly.

Of course it can be an argument for a proposed solution to the problem of universals that the solution would if true explain many things other proposed solutions do not explain. Thus explanatory considerations are far from irrelevant to the problem of universals, though they do not constitute the whole story. However, we must not confuse a question with the evidence

37 Swoyer is quite correct in pointing out that many arguments in metaphysics which have been conceived as demonstrative, such as the argument from design, would be far more plausible if understood as abductive arguments. However, there is no reason why this would have to hold of all ontological arguments, and I argue that it does not hold of most arguments for the existence of universals, which become circular if understood as abductive but may well be formally correct if understood as deductive.

38 People with radically different ontologies cannot even agree with regard to which rules of inference are valid, much less which premises are true. As I will show later in this dissertation, it makes a difference to the validity of many arguments for the existence of universals whether one uses a higher-order logic or a first-order predicate logic, as well as whether one uses a standard predicate logic or a free predicate logic.

39 Of course abductive arguments are no better in this respect. Opponents can deny the truth of the explanandum (if the explanandum is a sentence) or the existence of the explanandum (if the explanandum is a fact or some other entity) just as well as the truth of the premises of a demonstrative argument. When it comes to the validity of abductive inference matters are even worse; there is not even any precise theory, much less agreement, about what abductive arguments are good abductive arguments, i.e. what are abductively valid rules of inference (in the extended sense in which the word “valid” can be used here). So demonstrative arguments, though far from utterly certain, are on the whole a bit closer to being certain than abductive ones.
for its possible answers, or we are guilty of the verificationist fallacy. It seems to me that quite many ontologists who have taken the problem of universals to be a problem of explaining something have been guilty of this fallacy. Of course, if the explanatory power of a solution to the problem were the only possible evidence for its correctness, this distinction might not be of much importance. However, I will show that the explanatory power of a solution need not be the only evidence there is for the correctness of that solution.

Indeed, I will argue that at least when it comes to the first of the questions Porphyry asked, the explanatory power of a theory is not even relevant, since appealing to it would involve vicious circularity. However, I will not deny that when it comes to the other questions Porphyry asked, the remaining parts of the traditional problem of universals, explanatory power may be relevant. Of course the two additional questions Porphyry asks are not the only questions that an ontologist must ask if he concludes that there are universals, nor would they probably be thought to be the most important questions by a modern ontologist. Many other questions will be raised such as e. g. what universals there are and how many universals there are. If abductive reasoning is ever valid, contra van Fraassen et all, then it is probably a correct method to use in solving such questions, which constitute the remaining part of the problem of universals as it appears in contemporary discussion as a whole. In any case some kind of non-demonstrative reasoning, whether abductive or inductive, will be needed to solve at least some of these remaining questions. However, I will in the rest of this dissertation argue that such explanatory considerations are not relevant when it comes to the first, most fundamental part of the problem of universals, the question whether there are any.

6.4.1 Explanation, Description and Perception

Before explaining any data we need to describe them, as we can only judge if a proposed explanation is a good explanation by comparing the proposed explanans with the explanandum and seeing that the proposed explanans does explain the explanandum satisfactorily. It can be argued plausibly that different theories of explanation give different criteria for when a proposed explanation is a satisfactory explanation. E. g. a deductive-nomological theory demands (very roughly) that the explanandum imply the explanans, a unificationist theory that it unify the explanantia, a counterfactual theory that there be a true counterfactual in which the negation of the explanans is the antecedent and the negation of the explanandum the
we already need to posit universals or tropes or classes in order to correctly
describe any data we have. The solution of the Problem of Universals
involves not only what we must postulate to explain some phenomena but
more fundamentally what we must assume in order to describe these phe-
nomena properly, i.e. it brings with it problems concerning what is it that
given to us in experience.

The universalist and the trope theorist and the class nominalist and the
extreme nominalist often differ not only in what they think is needed to
explain some data. They differ already in how they would describe any
data that are to be explained; indeed, their biggest and most important
differences are over this question. Most extreme nominalists, of course,
think we only perceive concrete individuals. However, many universalists
(e.g. Gustav Bergmann and Reinhardt Grossmann) think that we also
directly perceive universals. For example in [Ber57, page 332], Bergmann
says:

I am directly acquainted with such things as, e.g., sensa and
some of the characters they exemplify.

consequent, and so on. I do not have to take any position here with regard to the ques-
tion which of these theories of explanation is the correct one. In any case we cannot say
whether any of these desiderata holds if we are not capable of describing what is to be
explained, of telling what the explanandum is.

According to standard theories of explanation (see e.g. (R4) in [HO48, §3]), the
explanans consists of true sentences. If we think there are facts, we can also alternatively
take the explanans to consist of the facts which such sentences correspond to either in
the sense that the sentences denote these facts or in the sense that these facts make such
sentences true. However, if our description of what our data are is incorrect, then the
description is not a true sentence but a false one and there is no fact corresponding to it,
and thus neither it nor any fact corresponding to it can be a constituent in a satisfactory
explanation.

Rodriguez-Pereyra admits that philosophers are not clear about what a solution to the
problem of universals should explain. However, he does not see that if the problem were
indeed an explanatory problem, then if different philosophers had different explananda,
they would be dealing with different problems, not one common problem as Rodriguez-
Pereyra presumes. Probably he thinks that though philosophers are not clear about
what the problem is, they yet have confusedly in mind the same problem, and that a
problem about how to explain something, though their unsuccessful attempts to explicate
it result in different problems. However, he presents no evidence for such a claim. It seems
to me that the most the philosophers addressing the problem share is a common question
about what there is; they do not share any common belief that could serve as a common
explanandum. Indeed, it could even be questioned whether they have a common question
in mind, since different philosophers have different concepts of universals; however, I have
already suggested that we can find a common concept of a universal which is contained in
all the more specific concepts of universals, and if this is correct then we can indeed find
also a common question presupposed by more specific questions.
These characters are of course universals and the exemplification here spoken of is the same thing as predication in Aristotle’s definition of universals, if that is taken to be metaphysical predication i. e. a relation between non-linguistic entities. Also many students of Bergmann argued in favour of the same view, e. g. Edwin B. Allaire in [All63].

Unfortunately, the value of the testimony of Bergman and his students is lessened by the fact that their view was they were yet influenced by verificationism and their view was motivated by their belief that we would have to be acquainted with any ontological simples in order to know at all that they existed43. Another student of Bergmann, Herbert Hochberg attacked their view in [Hoc65].

Hochberg held that the existence of entities such as universals and bare particulars is known on dialectical grounds, i. e. by means of inference; unfortunately, he did not make it fully clear whether such inference is supposed to be demonstrative or non-demonstrative, and if the later whether inductive or abductive. However, I think his words suggest a demonstrative inference; he speaks of an ontology providing a satisfactory basis for the analysis of ontological problems. Therefore his view does not support the view of such philosophers as Rodriguez-Pereyra and Swoyer that universals or tropes would have to be known by means of abduction.

In any case, the reasons Hochberg gave against Bergmann’s and Allaire’s view are not very convincing. Hochberg said in [Hoc65, page 124] that some may believe that what is presented (which meant in his terminology the same as what we are acquainted with) must be independent or distinguishable. This apparently meant for him that it must be capable of being presented separately, and he held that we could not be presented with universals without being presented with particulars or conversely, and this was a reason to think we were not presented with them at all. However, I see no reason to think that what we are acquainted with should have to be capable of being presented separately.

This view that we can be directly aware of universals can take two forms.

---

43 Allaire distinguished two uses of the word "know" in [All63, page 6]. In one use to know something is to be acquainted with it and in the second use of the word to know something means to recognize it. Allaire held that we know universals in both senses of the word, while we know bare particulars only in the first sense of the word. Curiously, Allaire entirely forgot the sense of the word in which to know something is to know it by description. Did he mean to implicitly deny that the word had such a sense?
It can be understood so that we are acquainted with universals with some sort of non-sensory, intellectual intuition. This seems to have been the view of Plato (and Platonists) with regard to his Forms (if they can be understood as universals). This may have been the view of Husserl and other phenomenologists in their famous doctrine of eidetic intuition (Wesensschau). Such a special kind of intuition has understandably been regarded with suspicion, though even such great logicians and mathematicians as Kurt Gödel have spoken on behalf of it (e. g. in [Gö1] Gödel defended a phenomenological method in mathematics). Nevertheless an appeal to such an intuition might furnish a rather weak argument for the existence of universals if it were the only reason to believe in them.

However, the view that we are directly aware of universals has also been held in the form that we perceive universals with our senses, in the same way in which we perceive sensible particulars. This view may be quite ancient and have multicultural respectability. Chakrabarti argues in [Cha75, page 367] that already the ancient Indian Nyāya-Vaiśeṣika philosophers held that most universals are known through sense perception. Whether this is correct or not, which I am not qualified to determine, most recent realist ontologists who have thought that we can perceive universals have in any case held such a form of the view. Grossmann has perhaps stressed this view most in [Gro90] as part of an overarching project to combine empiricism with realism (what he calls his fourth way)44. Grossmann says in [Gro90, page 131]:

> Perception acquaints us not only with individual things, but also with properties. It acquaints us not only with spatio-temporal entities, but also non-spatial and atemporal things. The senses

44There is a very widespread view that empiricism is somehow incompatible with realism concerning universals. E. g. Carnap said in [Car50] that

Empiricists are in general suspicious with respect to any kind of abstract entities like properties, classes, relations, numbers, propositions, etc.

However, though widespread, the view lacks any rational ground. It seems to me that there is no reason why abstract entities would present any special problem for any empiricism moderate enough to be compatible with any kind of realism, even realism concerning sensible particulars. Of course, at least most forms of Carnap’s empiricism were so strong that they were not compatible with realism even regarding concrete particulars. Carnap tried to reconcile empiricism with the use of abstract entities with the aid of his ontological relativism; however, I have already argued that this relativism does not work. Philosophers such as Bergmann, Grossmann, Fales and Armstrong present a far better way of reconciling empiricism with realism.
are a window through which we see not only the realm of be-
coming, but also the realm of being. The senses put us in touch
not only with changing things, but also with timeless universals.

Grossmann is far from the last to try to combine a strong form of em-
piricism with a strong form of realism concerning universals. Evan Fales
is a later representative of this view in [Fal90]; see [Fal90, pages 168-170]
for a summary of his view. Fales’s form of realism is stronger than that of
Armstrong and like that of Grossmann, since he admits that it is a form of
Platonism; according to him [Fal90, page 190] universals are not in them-
selves in space and time. Fales’s empiricism, however, is rather too strong in
my view, amounting (as he admits) to a form of neo-verificationism, which
I argue against in Section 5.2.1 of this dissertation.

While Russell and Bergmann had held that we can only be directly
aware of the characteristics of sense-data (aka sensa aka sense-contents),
Grossmann held that we also can be directly aware of the characteristics
of physical bodies. This seems to be the most common form in which the
view that we are directly aware of universals is held by philosophers today.
However, it is not unanimous, but Fales has returned to the older view of
Russell and Bergmann that we are only directly aware of the properties
of sense-data, as the only particulars of which we are immediately aware
are also in his view just sense-data (see e. g. [Fal90, page 258]). I will
not discuss in this work the question whether the properties we perceive
are physical properties or properties of sense-data (or of mental acts); in
order to support my view that realism concerning universals can be justified
without abduction it is enough to argue that we are directly aware of some
universals.

Grossmann also thought that we can sensorily perceive other abstract
(in the sense of non-spaitiotemporal) entities such as classes and numbers
(which he did not identify with universals or reduce to universals as I would
be inclined to do). Of course Grossmann’s claim may be more plausible with
respect to such obviously sensory properties as shapes and colours than with
respect to such more abstract entities as numbers.\footnote{It is not clear that Grossmann’s attempt to combine empiricism with realism succeeds
fully. Even if we can perceive universals and their relations, and even if we can perceive
such highly abstract universals or other abstract objects as numbers and classes, we cannot
perceive transfinite numbers or classes. Grossmann’s theory implies that we could perceive
transfinite numbers or classes only if we perceived that there are infinitely many entities

458
The view that we perceive universals with our senses is often, as in Grossmann, part of the view that we perceive facts or states of affairs. According to this combined view we perceive a state of affairs consisting of particulars and universals and we perceive it as consisting of them. In a way even Armstrong agrees with this view, though there is caveat to be made. Armstrong subscribes to a propositional theory of perception in [Arm97, §7.11, page 96]. This propositional theory of perception contrasts with the view of many epistemological coherentists and other opponents of foundationalism that perception cannot justify any beliefs because it is not propositional (so suggesting that Armstrong’s theory is a foundationalist or foundherentist one). E. g. Donald Davidson said when defending a coherentist theory in [Dav86, page 311] that the relation between a sensation (which Davidson does not distinguish from a perception) and a belief cannot be logical, since sensations are not beliefs or other propositional attitudes. Davidson’s argument might not work even if we admitted his premise that perceptions are not propositional, since as Susan Haack points out in criticizing Davidson’s theory in [Haa09, page 112], justification is not purely a logical matter. However, the premise that perceptions are not propositional can also be questioned (even though Haack seems to accept it), and Armstrong denies it (without referring to Davidson). A propositional theory of perception suggests that perceptions are propositional attitudes and therefore their contents can be in a logical relation to perceptions, contrary to Davidson’s view.

Perception has also been viewed as propositional by philosophers of some kind: however, it is rather obvious that we cannot perceive such a thing. This is indeed not wholly uncontroversial; there have been philosophers who seem to have denied this; e. g. Leibniz seems to have held that we have subconscious perceptions of all the infinite number of monads there are. However, most philosophers today find this quite implausible, and in any case if such perceptions are subconscious, it is not clear if they could yield conscious knowledge of mathematical axioms. So Grossmann’s theory still leaves it quite obscure how we can know that such mathematical axioms as the axiom of infinity are true. The problem that mathematical knowledge has traditionally presented for empiricists may therefore not be solved even by Grossmann’s variant of empiricism; even if we assume that we can perceive universals with our senses, we might still need a non-sensual kind of intuition in order to know that axioms such as the axiom of infinity is true, if this cannot be explained as a result of induction. However, I think that Grossmann’s view that we can perceive universals is independent of his empiricism and is very plausible whatever view we hold of his empiricist theory and at least suffices to justify some kind of realism concerning universals.

In [Arm61, pages 105-106] Armstrong held that perceptions can be analysed with the aid of the notion of belief - that perception is the acquiring of an inclination to believe in particular facts about the physical world by means of our sense - which of course implies that they have propositional contents. However, it is quite possible to hold a proposi-
not primarily motivated by epistemological or metaphysical motives, as by many philosophical logicians - as Ilkka Niiniluoto says in [Nii82, page 116], Jaakko Hintikka and his followers treat perception as a propositional attitude. This surely serves to confirm the propositional theory of perception. In fact Armstrong holds that the objects of beliefs are states of affairs or facts, not propositions, so the name ”propositional theory of perception” is a bit misleading. One might think that this would give rise to the same problem about how perceptions justify beliefs as Davidson has. However, for each state of affairs there of course corresponds a proposition according to Armstrong’s theory, the strongest proposition it makes true, so there can be a logical relation between that proposition and the contents of the belief states it justifies.

However, there is one important caveat to be made in ascribing to Armstrong the view that universals are perceived. In Armstrong’s view the states of affairs we perceive do not contain as constituents genuine universals (as Grossmann thought) but what he calls second-class properties and relations (which many ontologists would following David Lewis call abundant properties and relations together with Armstrong’s third-class properties and universals), which in his view supervene upon genuine universals (which Lewis would call sparse properties and relations). It naturally follows from this that the states of affairs that are perceived are also second-class states of affairs.

However, these results come about because Armstrong’s conception of a universal and a state of affairs is very narrow (though not in a precise way); other ontologists would call these second-class properties and relations universals, since they clearly are not tropes but many of them are instantiated by many particulars. It is surely more plausible that if there is an objec-

---

460
tive distinction between sparse or first class and abundant or second and third class properties and relations, then only the latter can be perceived, while the former have to be inferred inductively or abductively (just as any fundamental particles or fundamental physical events would have to be). If Grossmann thought that sparse properties and relations could themselves be perceived, as many of his statements suggest, then his view is less plausible than Armstrong’s.

In any case, though Armstrong may not always be consistent in this, he yet claims explicitly that these second-class properties and relations really exist. He says in [Arm97, §3.9, page 45] that while the second-class properties are not properties additional to the first class properties, yet it is to be emphasized that this does not make the second-class properties unreal. They are real and cannot be talked away.

This is already a substantial and an extremely controversial ontological claim. Extreme nominalists and most trope theorists would deny that such second-class properties and relations exist and that they can be perceived. It seems to me that the crucial, primary disagreement between Armstrong and many other ontologists such as extreme nominalists or eliminative trope theorists is whether these second-class properties and relations exist and can be perceived, or whether it is rather quasi-tropes that are perceived (as Campbell would say). I would say that agreeing that these second-class properties and relations exist would already constitute a partial solution to the problem of universals; it already commits Armstrong to realism about entities which

47 Opponents of realism about universals have seldom addressed this question. One interesting recent attempt to criticize the view that states of affairs can be perceived is Mark Textor’s [Tex09]. Textor perspicuously draws attention to the phenomenological questions and questions concerning the logical analysis of perceptual statements that I hold are crucial in solving the basic part of the problem of universals. However, Textor commits a serious mistake already in the title of his article. He purports to criticize Armstrong’s theory among others yet he contrasts particulars with states of affairs and asks whether particulars or states of affairs are given in perception, clearly presupposing that the two disjuncts are incompatible. However, according to Armstrong (see [Arm97, §8.4, pages 126-127]), states of affairs are particulars (though they contain universals as their constituents), so the only answer Armstrong could give to Textor’s question would be “Yes.” I do not think that this is the kind of answer that Textor would have anticipated! Textor’s criticism might (though there are also other problems with it that I cannot deal with here) work against a theory that took states of affairs to be such as Chisholm or Plantinga think and held that these states of affairs would be objects of sensory perception; however, it is not clear whether any philosopher has held such a view, so this criticism may well be directed at a pure straw man. Perhaps it might be more effective against Grossmann’s theory than Armstrong’s, but I am not sure even about this.
are not particulars. Also since Armstrong holds that we can directly perceive these second class properties surely no abductive argument is needed in order to come to the belief that they exist. I am willing to admit that the inference from second-class properties and relations to the sparse universals upon which they supervene may be abductive; it is plausible in any case (though not obvious) that it is non-demonstrative. Whether it is a good or valid (in the extended sense in which a non-demonstrative argument may be valid) argument is a different question that I cannot address here. However, there is little point in arguing over whether this non-demonstrative argument from the second-class properties and relations to sparse universals is valid if there is no agreement about whether these second-class properties and relations exist and so no agreement about whether the premises of that non-demonstrative argument would be true. Therefore the question of the validity of that inference is a question of lesser importance.

This view fits well with a causal theory of perception if it is combined with the highly popular theory of causation according to which facts or states of affairs are the members of the causal relation (in the most fundamental sense of the word “causal”). Most of its adherents do naturally hold some kind of causal theory of perception. Grossmann and Fales indeed reject attempts to reduce perception to causation. Grossmann says in [Gro90, page 250]:

When you see that there are two apples on the table in front of you, your mental act of seeing stands in a unique, noncausal, relationship to a certain fact.

Fales similarly says in [Fal90, page 259] that as far as meaning goes, the essential feature of perception is intentionality and the fact that it involves

48It is not clear that Armstrong would agree with this view. He thinks that the particularity of particulars is part of the content of perception, but nevertheless, he stresses that he does not claim that questions of individuation can be settled by a propositional theory of perception plus perfunctory phenomenological inspection. One might infer that he would then have to hold that analogously the question of the existence of second class properties is not to be settled by perfunctory phenomenological inspection either, but it is not clear what his justification for this would be. What about careful phenomenological inspection? At the very least even if we are only weak foundationalists, if second class properties and thin particulars are part of the content of perception, this should give prima facie justification to the claim that they exist, and not just show that it is not obvious that they do not, which is all that Armstrong claims it does. For an interesting correspondence between Armstrong and Grossmann that brings out their differences in metaontology see [CT09].
some kind of sensuous consciousness. However, also Grossmann admits in [Gro90, pager 240] that if one perceives, in the literal sense, a certain fact, then that fact must play a causal role in one’s perceiving it. Fales also says in [Fal90, page 259] that when one perceives something which is not a sense-datum, it is necessary that the object perceived stand in some suitable causal relation to the act of perceiving it. Even Grossmann and Fales then admit that there is an intimate link between perception and causation, even if it is not sufficient for a reduction of the former to the latter. Armstrong, however, might be in favour of a reductive causal theory of perception.

However, the view that we perceive universals with our senses can also be combined with different theories of individuation (and perception). A philosopher might also hold that particulars are complexes of universals and that in perceiving such a complex we often perceive some of the universals of which it consists (but do not perceive the particularity of particulars as Armstrong thinks). This view had already been held before Grossmann, for example by the neo-realist William Pepperell Montague (do not confuse him with the more famous Richard Montague) in [Mon25]. Montague says in [Mon25, page 77]:

Both rationalists and empiricists commit the same error, for they regard the originally given elements in experience as particulars and nothing but particulars, and as lacking anything that can be called universality. ... Experience is indeed originally of particulars, that is, of objects that are presented at particular times and places. But each of these experienced objects has a universal nature which is as indefeasibly its inclusive property as is its unique position in space and time its exclusive property. In other words, the given elements of experience are complexes of universals, each complex being associated with a particular position in the space and time series. It is this latter factor of position which constitutes particularity and makes each individual numerically different from every other individual. To form the concept of a universal, it is, as we have already seen, not necessary for the mind to manufacture or create anything different from what is given, but only to abstract the attention from the particular position of the given complex and concentrate it upon
some one or more of its qualitative or non-positional elements.

Montague’s view differs from Grossmann’s as Montague is committed to a kind of bundle theory of particulars, which Grossmann opposes, since Grossmann favours a theory of bare particulars as individuating principles. I think that a bundle theory is in fact better, so Montague seems to me to be more correct in this respect49; however, what ontological position we take with regard to this highly difficult and advanced matter does not affect the more basic epistemological point that universals can be perceived when they are ingredients of particulars.

Many trope theorists on the other hand think we do not perceive universals, but perceive tropes directly50; e. g. Kevin Mulligan, Barry Smith and Peter Simons give in [MSS84, §4, page 304] as a general defence for the existence of moments (which is their word for tropes) that they can be objects of perception and John Bacon goes so far in [Bac95, page 4] as to argue that tropes are epistemically prior to both universals and concrete particulars.

This makes both universals (in the sense of the word that contains second-class and third-class properties and relations) and tropes very different from such entities as electrons or black holes, which are paradigmatic examples of entities whose existence is inferred abductively; everyone who agrees that there is any clear distinction between observable and unobservable entities thinks that electrons and black holes are unobservable, while as we saw many philosophers think that universals or tropes are observable and indeed frequently observed. Some entities whose existence is plausibly inferred abductively are indeed observable according to most conceptions

49There are of course well-known difficulties with bundle theories, and with the kind of theory Montague sketches in particular. Montague seems to be saying that experienced objects are individuated by their spatio-temporal positions; however, this raises the question of what these positions are ontologically and how they are individuated. If such positions are taken to be particulars in accordance with an absolutist conception of space-time, then it can be argued that Montague’s theory just pushes back the question of the individuation of particulars.

50Interestingly at least some pragmatists seem to agree (at least at some times) with this view. Though Peirce, the founder of pragmatism, had apparently argued against all kinds of foundationalism in [Pei68], Williams James supported in his radical empiricist period in [Jam22, page 42] a form of empiricist foundationalism according to which some relations, namely relations between experiences (understood by James in a neutral monist sense as in themselves neither mental nor physical), are themselves experienced. Since James holds in the same work (in [Jam22, page 41]) that universals are mere abstractions, this seems to imply that these experienced relations are tropes (though James does not give any clear or explicit statement of his view with regard to the problem of universals).
of observability, such as the planets Mercury or Pluto, whose existence was first postulated to explain perturbances in the orbits of other planets. However, even in their case it is uncontroversial that they either are in fact unobserved or at least were unobserved when they were first postulated on the basis of abduction. In this they differ from (abundant) universals and (quasi-)tropes, which according to many philosophers who postulate them have always been quite commonly observed.

Any argument for their existence would then not be an abductive or inductive non-demonstrative argument but a demonstrative argument that only tries to draw attention to what anyone already implicitly knows. That the arguments for the existence of universals or tropes must be like this has already been seen by many philosophers, e. g. by Nicholas Wolterstorff in [Wol70, page xiii]. Wolterstorff indeed holds that all ontology is descriptive; while I would agree that ontology is primarily descriptive, I would not go so far as to deny that it can sometimes be explanatory in a non-trivial sense, if the word is used as broadly as it is often used today.

At least a direct realist about material objects (like Grossmann or Armstrong or Pols) would say that we do not need any abductive argument to show us that there are material objects. In their view we might indeed need an abductive argument to show of some material objects that they exist, e. g. of electrons, but not to show us that there are some other material objects, e. g. stones or trees. A philosopher who thinks that we are directly aware only of sense-data (like Russell, Bergmann or Fales) might say that the existence of all material objects is inferred abductively to explain our sensations or sense-data (unless he is a phenomenalist or neutral monist, in which case he requires no abductive or inductively inference, since he thinks he can construct material objects logically out of sense-data). However, he could not then say that we need to infer our sensations and sense-data abductively from anything else. Similarly if the universalists or trope theorists mentioned above are correct, then we do not need any abductive argument to show that there are universals (or at least second-class properties and relations) or (quasi-)tropes either.

At best such an argument could then strengthen the justification of an already justified belief. However, even this would require that the explainandum serving as the premise of the abductive argument would be independent of the belief that it would further justify; if it is not so, then such an argu-
ment is wholly circular, just like postulating that there are material objects in order to explain the existence of material objects would be. I will argue later in this dissertation that at least most abductive arguments that have been presented for the existence of either universals or tropes are circular in exactly this way. Therefore before trying to find any abductive argument for a solution to the problem of universals one should examine whether these philosophers who think that universals or second-class properties and relations or tropes or both can be perceived directly are correct. Nor is this an easy task; rather solving it will form the major part of an argument for any solution for the problem of universals.

Rodriguez-Pereyra fails to do this. He argues in \([\text{RP02, §12. 6, page 210}]\) that the decisive superiority of resemblance nominalism over universalism and trope theories consists in its following Russell’s supreme maxim of scientific philosophizing (as expressed for example in \([\text{Rus18c, page 155}]\)):

> Whenever possible, logical constructions are to be substituted for inferred entities.

, or as Rodriguez-Pereyra reformulates this, to avoid postulating ad hoc entities. By the ad hoc ontology of a theory Rodriguez-Pereyra understands those entities postulated by the theory the only or main reason to believe in which is that they contribute to the explanation of some phenomenon. I am not at all sure that Russell’s maxim is a correct maxim; it is a predecessor of logical positivism’s verificationism\(^{51}\) and has generally been used to support anti-realistic theories, and today many philosophers, especially most scientific realists, would deny the correctness of the maxim\(^{52}\). However, even if it

---

\(^{51}\)Russell’s maxim must be sharply distinguished from Russell’s earlier Principle of Acquaintance, which as I showed was far weaker than verificationism. The Principle of Acquaintance does not imply the maxim; it does not in the least dissuade us from inferring the existence of entities with which we are not acquainted, as we can have have knowledge by description of them. The principle said nothing about whether we can infer that some entities exist, since it did not concern propositional knowledge at all but only knowledge of things, and inference is a method of reaching propositional knowledge. The principle and the maxim come from two different stages in Russell’s philosophical development, and Russell’s adoption of the maxim led to a big change in his ontological views, away from dualism to neutral monism. This remark is useful, as some philosophers do not keep the principle and the maxim apart clearly enough; for example, Misak seems to confuse them in \([\text{Mis95, page 44}]\). By conflating the principle and the maxim, Misak produces an excessively positivistic image of Russell’s philosophy.

\(^{52}\)The maxim led Russell into phenomenalism and neutral realism since he thought that we were directly aware only of sense-data. A direct realist might avoid this consequence, and admit that some physical objects, namely observable, macroscopic ones, are known
were valid, I do not think it can be legitimately used to support resemblance nominalism in the way Rodriguez-Pereyra uses it.

It seems at first sight that the very point of Russell’s maxim is to deny the legitimacy of all abductive (and inductive) inference, and if this is so then it is hardly consistent to use the maxim and yet hold that the problem of universals should be solved by abductive inference, as Rodriguez-Pereyra does. If you completely avoid postulating ad hoc entities, then you are not using abductive inference. Indeed if the problem of universals were an explanatory problem, then someone consistently following Russell’s maxim might have to declare the whole problem a pseudo-problem. In fact it can be argued that Rodriguez-Pereyra does violate the maxim in postulating possibilia and in postulating classes as primitives, so he does use abductive inference. He would say that he does not postulate possibilia only to solve the problem of universals, but primarily to provide truth-conditions for modal discourse. Similarly he would say that he postulates classes primarily to account for mathematics. It is not clear that Russell would have thought that this was legitimate either; surely Russell’s maxim is not meant to apply only to the problem of universals but to any problem, including the problem of accounting for modal discourse and the problem of accounting for mathematics.

However, the maxim is better interpreted as not rejecting abductive and inductive inference completely but just as seeking to minimize their use. There are significant differences between Russell’s maxim and early ver-}

\[53\]In fact the solution of the problem of universals cannot be separated from the solution of these other problems, and could not be separated from them even if the problem of universals were an explanatory problem. In order to respond to an extreme nominalist who denies the existence of classes as well as that of intensional universals or to a universalist who denies the existence of classes but believes in intensional universals Rodriguez-Pereyra would have to appeal to facts involving modal discourse and mathematics, which his opponent would of course try to interpret in his own way. To solve such controversies we must surely ask which ontological theory can give the overall best solution to the whole complex of issues involving the problem of universals, the interpretation of modal discourse and the interpretation of mathematics. It is no good if one theory can give a better account of one set of data if this forces it to give a worse account of other, equally important data.

\[467\]
licationism. Russell’s maxim is a heuristic or methodological maxim, not a semantical claim about significance\textsuperscript{54}. Russell said that inferred entities should be avoided and replaced with constructions \textit{whenever possible}, but unlike early logical positivists, he did not suppose that it would be \textit{always possible}. Nor did Russell assert, like van Fraassen, that whenever a purportedly inferred entity could not be constructed out of entities with which we were acquainted, we would have to suspend judgement about its existence (though he was tempted to do so). Indeed, Russell’s maxim must be interpreted in this way if it is to be even compatible with others of Russell’s own views. Russell explicitly states himself in [Rus18c, page 157-158] that he is allows inferred entities of two kinds, the sense-data of other people and unsensed sensibilia, out of which he constructed physical objects or substitutes for them. Therefore Russell might have come to see the positing of possible particulars as no more problematic. Russell says that inferred entities should be similar to those which are given and Rodriguez-Pereyra argues that the introduction of possibilia involves a less radical departure from an ontology entirely populated by actual concrete particulars than the introduction of universals or tropes, because possibilia are according to him of a kind with actual concrete particulars. Presumably Rodriguez-Pereyra thinks that possibilia are this respect analogous to unsensed sensibilia, which Russell supposed to be of a kind of with sensed sensibilia i. e. sense-data, so that their introduction involved a lesser departure from an ontology entirely populated by sense-data that the introduction of physical objects as primitive entities. Rodriguez-Pereyra can quite reasonably argue that because of this what he does violates the spirit of Russell’s maxim less than universalists and trope theorists do.

However, I will show that it is questionable whether he really violates the spirit of the maxim less than his competitors - he may in fact violate it more than universalists or trope theorists. First of all, it is not clear what Rodriguez-Pereyra would say about classes. They are surely not of a kind

\textsuperscript{54}Richardson seems then to be radically wrong when he claims in [Ric90, page 7] that

Hence, the supreme maxim subserves an ontological and analytic fact: entities we are acquainted with are all the entities there are

This is not quite clearly formulated, but Richardson seems to be saying that according to Russell it is an analytic truth that the entities we are acquainted with are all the entities there are. If this is what Richardson meant, he is completely wrong; Russell did not even think that this is a truth, and if it was it would surely according to him be a synthetic truth.
with concrete particulars. His postulating of classes as primitives can thus surely not be considered as analogous to Russell’s postulation of unsensed sensibilia, and it seems it is thus in any case in flagrant contradiction with Russell’s maxim. It does not help much to say that the introduction of classes is done for reasons independent of the problem of universals; if violating Russell’s maxim flagrantly is acceptable in solving other problems, why would the problem of universals be an exception? Russell himself tried at a time to reduce classes to propositional functions, which he then held to be universals, but Rodriguez-Pereyra as a nominalist naturally cannot do this. The only way Rodriguez-Pereyra could escape this problem would be to maintain that we are directly aware of classes. I am not sure if he would be willing to do this, though it is not at all an absurd position, but many competent philosophers (e.g. Reinhardt Grossmann and Penelope Maddy in [Mad80] have held it, although in different forms\textsuperscript{55}. Secondy, Rodriguez-Pereyra assumes that universals or tropes would be inferred entities, so that assuming either of them would be a disadvantage for universalism and trope theory. However, if Grossmann or Bergmann are right, then universals are perceived and hence we are directly aware of them and hence they are not inferred entities in Russell’s sense. In fact Russell himself thought (at least in many phases of his ever-changing views) that we are acquainted with universals\textsuperscript{56}, so they would not be ad hoc entities in his view. On the other hand if Bacon or Mulligan, Simons and Smith are right then tropes are not inferred entities in Russell’s sense either.

Thus universalism and trope theory need not have any bigger ad hoc ontology than resemblance nominalism if these philosophers are right. Therefore Rodriguez-Pereyra must show that they are wrong if he is to be justified for claiming that resemblance nominalism is the superior theory because of

\textsuperscript{55}Maddy held that only concrete entities can be perceived and that classes are concrete while Grossmann held (see \cite{Gro90}, pages 240-245) that classes are not concrete (i.e. spatio-temporal) but that abstract entities can be perceived if they are ingredients of concrete entities.

\textsuperscript{56}E.g. Russell says in \cite{Rus12b, page 158}:

\begin{quote}
It is obvious, to begin with, that we are acquainted with such universals as white, red, black, sweet, sour, loud, hard, etc. i.e. with qualities which are experienced in sense data.
\end{quote}

Russell continues by saying that we are also acquainted with relations. Most obviously there are relations between different parts of a complex sense-datum, but there are also higher-order relations between universals themselves. Indeed, Russell tries to found a priori knowledge in knowledge of relations between universals.
best avoiding ad hoc ontology. He does indeed attempt this in the case of tropes, but not in the case of universals\textsuperscript{57}

Even in the case of tropes his argument is not very strong. Rodriguez-Pereyra says that the fact that we perceive tropes is an independent reason to believe in tropes only if it is clear that we perceive tropes. Surely this is not correct; if it is not clear whether we perceive tropes, then all that follows from this is that it is not clear whether our perceiving tropes is an independent reason to believe in tropes and thus that it is not clear whether tropes are ad hoc entities. In such a situation it is not clear that any theory would be superior to another because it avoids postulating ad hoc entities either. Rodriguez-Pereyra tries in \cite{RP02, §4.12, pages 93-95} just to show that the argument of Mulligan, Simons and Smith for the conclusion that we perceive tropes is not valid. Even if the argument for this were correct, it would not amount to showing that resemblance nominalists postulate less ad hoc entities than trope theorists do. Trope theorists like Bacon and Campbell who think that tropes are not only ontologically, but also epistemologically primary entities can quite consistently hold that concrete particulars are inferred entities in Russell’s sense, entities postulated to explain the similarities and regularities we observe among tropes. They can thus hold that in accepting concrete particulars as ontologically basic it is resemblance nominalists who are violating Russell’s maxim. If it is not clear whether we perceive tropes, it is not clear whether we perceive concrete individuals either. Just as Rodriguez-Pereyra tries to explain away apparent perception of tropes, so trope theorists can try to explain away apparent acquaintance with concrete individuals. They might argue that even if we in some loose sense perceive concrete individuals, yet such perception contains inference as its part and is not direct awareness of its objects, so the objects of such perception are inferred objects in Russell’s sense (as Russell himself thought was the case when it came to the perception of physical objects, since according to him we were directly aware only of sense-data and not

\textsuperscript{57}He does indeed say in \cite{RP02, page 93} that he agrees with Mulligan et all that it would be counterintuitive to say that we see things as exemplifying certain universals. However, this scarcely amounts to an argument; it is all too easy to settle any controversial question by just insisting that all theories except one’s own are counterintuitive. Of course, it might turn out that no genuine argument can be found, so that philosophers with different intuitions just cannot ever come to a rational agreement. However we should not accept this defeatist conclusion before exerting the utmost effort to find some way to rational agreement.
of perceived physical objects). Rodriguez-Pereyra argues that a sign of universals and tropes being a part of ad hoc ontology is that all parties to the dispute agree that concrete particulars exist, even those that, like bundle theories, reduce them to other entities, but not all parties accept universals or tropes. However, it is not clear that all parties to the dispute would accept that concrete particulars exist. Keith Campbell’s claims in [Cam90] seem to me to imply that there really exist only tropes, though Campbell is rather coy in saying this explicitly. Campbell holds that only tropes are ontologically basic, and holds further that only basic tropes really exist; even non-basic tropes or quasi-tropes like the colour of an object, are according to [Cam90, §6.11] just well-founded appearances\(^{58}\). Surely concrete particulars, being also derived from basic tropes (and in the case of ordinary particulars indeed only through the quasi-tropes) must then in Campbell’s system with a still better reason be mere well-founded appearances. Thus Campbell implies that concrete objects do not really exist but only appear to exist. It is no doubt the case that the existence of concrete particulars is less controversial than that of tropes, but this is surely irrelevant. Russell’s maxim does not mean that we should accept only entities which most people agree to exist, but that we should accept only entities we are acquainted with. We can be acquainted with entities whose existence most people deny because of confusion, and entities that most or even all people agree to exist can be entities believed in only because they help to explain something. Russell himself thought that entities like physical objects which most people would agree to exist are inferred entities, while more controversial entities like sense-data and the universals they instantiate were not inferred entities, and used to maxim to reject the first and accept the latter.

More importantly from a purely metaontological point of view, however, in discussing at all the question of whether we perceive tropes, Rodriguez-

---

\(^{58}\)The position of these quasi-tropes in Campbell’s system is similar to that of the second class and third class properties and relations in Armstrong’s system; both philosophers are somewhat equivocal about the ontological status of the respective entities in question. However, Armstrong never goes so far as to say that the second class properties and relations are mere appearances, even well-founded appearances, and implicitly denies this possibility in [Arm97, §3.9, page 45]; Campbell is then a more strictly fundamentalist and eliminativist philosopher than Armstrong (or other trope theorists such as Bacon). This may, indeed, constitute a big problem for Campbell when he tries to explain how we come to know that there are tropes; the argument will presumably be non-demonstrative, abductive or inductive, but neither inductive nor abductive arguments can validly lead to denying the existence of entities asserted to exist in the premises of these arguments.
Pereyra is discussing a question that is not a question of how to explain something but a question of what there is to be explained, and so is going outside what he considers to be the problem of universals. Thus he shows that his definition of the problem is too narrow for even him to follow in practice. However, his definition forces him to consider as only an incidental question what is really one of the most fundamental sub-problems of the problem of universals, the point from which really rigorous research should start.

Of course, those who argue that the existence of universals or tropes is already implicit in what anyone knows do not claim that ordinary people would explicitly know that there are universals or tropes or classes, but they think that it will come as a surprise to them, as a character in Moliere’s play learned with surprise that he had been speaking prose all his life.

The question whether we perceive universals can of course be formulated in what Carnap called the "formal mode", if someone yet thinks that this is required to make the question clear. Using semantic ascent and the results of other sections of this dissertation we arrive at the following formulation: Do protocol sentences i.e. observation sentences or the sentences of the metatheory in which the weakly observational terms which are the only non-logical terms which occur in the protocol sentences are interpreted include or logically imply sentences in which objectual existential quantifiers.

\[59\] Rodriguez-Pereyra says in [RP02, §4.12, page 95] that since Resemblance Nominalism is a metaphysical theory, no account of perception is required in his book. However, this is incorrect; since Rodriguez-Pereyra appeals to Russell’s maxim, this would require him to develop a theory of perception he does not develop. The very point of Russell’s maxim is just to make the solution of ontological questions dependent on the prior development of a detailed epistemological theory of perception, for without having a fully developed theory of perception we have no way of knowing which entities are inferred and which are not. In the absence of such a full theory there is no way to resolve the dispute between Rodriguez-Pereyra and Mulligan et al with regard to whether tropes are ad hoc entities or not. What Russell himself does in the works Rodriguez-Pereyra refers to is to develop a theory of perception and then draw ontological conclusions from it. Anyone who wants to honestly follow Russell’s maxim must do the same, though he can of course come to very different epistemological conclusions from those Russell himself came to. It may be argued that it is unintuitive that epistemology would be in this way prior to ontology, and I have some sympathy with this view. The argument based on ontological commitments which I will suggest for the existence of universals, though not abductive, will not depend on fundamental epistemological considerations either, but starts just from statements which most people would agree to be true, whether on the basis of abductive or inductive inference or on the basis of acquaintance, and tries to show that these statements imply the existence of universal. However, if one wants to hold consistently that epistemology of perception is not methodologically prior to ontology then he must deny the validity of Russell’s maxim.
fiers bind predicate variables or other variables that can be substituted for terms symbolizing predicate terms of natural languages? When Carl Hempel said in [Hem65, pages 102-103] that an observation sentence might be construed as a sentence which asserts or denies that a specific object or group of objects, of macroscopic size, has a particular observable characteristic, he implied that the answer to such question is affirmative and that we perceive properties, whether universals or tropes. Hempel might not have even noticed the important ontological commitments he incurs, so this may not be his considered view; however, this makes his testimony all the more important, as it shows that the assumption that we can observe universals is a very natural one to make, even inadvertently. In Hempel’s view the properties we perceive are those of physical objects, since he held that observation sentences belong to what Carnap has called the thing-language, and mentions that the phenomenalistic approach has many problems which he could not discuss at that place. However, Carnap also spoke of observable properties at the very beginning of [Car56, page 38], yet he held in [Car56, page 40] that it may suffice to use only individual variables in the observation language. While Carnap then speaks against the view the universals can be perceived, his support for that view obviously lacks conviction, as can be seen from the use of the auxiliary verb ”may”, which here clearly stands for epistemic possibility (and Carnap did not address at all the question whether individual variables suffice also in the metalanguage of the observation language). I will argue later that an adequate observation language has to use either predicate variables or set variables or other variables ranging over other entities than particulars, for the very simple reason that we do not have and cannot have names for all phenomenal properties. We may, for instance, want to say that two physical objects or sense-data have the exact same hue or shade of colour, but we do not have names for all hues or shades of colour and cannot have them, quite independently of whether colours are taken to be properties o sense-data or physical objects. If this is correct, it implies further that the predicate ”universal” (if understood as referring to abundant rather than sparse universals) is itself a higher-order observational predicate (more exactly, if we use the distinction between weakly and strongly observational predicates I defined earlier, a weakly observational predicate). The property of being an abundant universal (if there is one) is then an observable property, more exactly a weakly observable property.
Scientific Explanation and the Problem of Universals

So far I have shown that the problem of universals is not solely about explanation; now I will argue that it is probably not at all about explanation, at least not of the same kind of explanation as ordinary scientific explanation. I will show there are reasons to suspect that the whole problem introduced by Armstrong is a pseudo-problem since the kind of explanation he seeks for would be a pseudo-explanation. However, I will not be arguing that all explanatory problems that have been claimed to be identical with the problem of universals would be pseudo-problems. For instance, the Problem of Concepts that Gregory Salmieri formulated is probably a genuinely explanatory problem and is probably not a pseudo-problem; however, it is yet not identical with the Problem of Universals (though it is a closely connected problem) and is not even an ontological problem at all.

If the solution of the problem were of the same kind as an ordinary scientific explanation, then we would except that all or at least most of the proposed solutions should fulfil the criteria commonly demanded of a satisfactory scientific explanation. It is very strange that of all the many philosophers that have claimed that the problem of universals is an explanatory problem comparable to problems of explanation in science, almost none has tried to compare the treatment of the problem in any detail with the general theories of explanation proposed in philosophy of science. This suggests that their claim that the problem is concerned with explanation is a hollow claim. I cannot in this dissertation compare the theories of scientific explanation in any ultimately sufficient detail either. That would obviously be a tremendous task, since there are so many theories of scientific explanation and most are highly complex⁶⁰; however, I try to at least get the task of comparison to a good start and show that the prospects for the metaontological claim that the problem of universals is an explanatory problem are not promising. I will argue that if a solution to the problem of universals, at least to its most basic part, were an ordinary scientific explanation, it would not fulfil the criteria demanded of a satisfactory scientific explanation and would be a bad explanation.

Of course most theories of explanation are not about explaining how something can be the case but about explaining why something is the case.

⁶⁰For a good recent overview of such theories see [Woo03].
Perhaps then we should not except them to apply to explanations of the kind Rodriguez-Pereyra and Swoyer refer to. However, in this case we may not have any clear idea what an explanation of how something can be the case would involve. While scientific explanation of why something is the case has been studied at length, no agreement has emerged and every theory has difficulties. However, at least there are detailed theories. There are not even theories (at least not many systematic ones as in the case of why-explanations) about what it would be to explain how something is possible. Also, probably all examples of abductive inference outside ontology are inferences to the best explanation of why something is the case. This raises doubts whether the kind of reasoning supposed to be used in ontology can be abductive in the same sense as abductive reasoning outside ontology.

It is commonly held that the explanandum should not imply the explanans logically or analytically. This is indeed often only presupposed and not stated explicitly, as it is taken to be obvious; however, one of the classical philosophers of science, Ernst Nagel, has stated it explicitly. Nagel’s theory of explanation is a version of the Deductive-Nomological Model of explanation, as most famously presented in [HO48]. According to this model of explanation an explanation is an argument in which the explanans consists of premises which imply the explanandum (or, as Nagel calls it, the explicandum). This theory of explanation has been often attacked recently, and I would myself agree that it needs some modification if not outright rejection. However, though the critics object to many features of the Deductive-Nomological Model, I do not think that they object to this extremely intuitive idea that the explanandum should not imply the explanans. Nagel lists in [Nag61, page 34] criteria of a satisfactory explanation; one of them

\[61\] Nagel uses the word "explicandum" so that it is synonymous with the today more usually employed word "explanandum"; he does not mean the object of explication in Carnap’s more famous sense, which is closer to conceptual analysis than to scientific explanation.

\[62\] Other theories of scientific explanation typically fit the problem of universals even worse. E. g. the Causal Mechanical Model of explanation says that the explanans and explanandum are or describe events which are connected by a continuous causal process. This theory of explanation at least does not fit any proposed solution to the problem of universals. It is not a general theory of explanation, but only a theory of the explanation of events; however the explanandum of the problem of universals is conceived by those who think it to be an explanatory problem, it is surely not an event. That many things share a common property is surely not always an event, but is often something permanent or even timeless.

\[63\] Nagel is talking about the explanation of laws, which does not quite fit the kind of
is that

the premises, taken singly or conjointly, do not follow logically from the explicandum.

Nagel says also in \cite{Nag61, page 38} that

we expect the explanatory premises in a satisfactory explanation to assert something more than what is asserted by the explicandum.

Nagel uses here a broad concept of logical consequence that includes analytical entailments, as can be seen from the fact that he gives as an example of a pseudo-explanation that does not satisfy this requirement the explanation in Moliere’s famous play of the fact that opium induces sleep with the dictum that opium possesses a dormitive virtue.\footnote{Nagel’s criteria thus presuppose a distinction between analytical and non-analytical statements. Some may therefore object that they should be rejected because Quine has in their view shown this distinction to be untenable. I would not agree that Quine has shown anything like that, but this dissertation is not a place where I could defend such a claim. However, my main point would hold even if Quineans were right in rejecting the distinction between analytical and other truths. Surely the thesis of Rodriguez-Pereyra, Swoyer and others which I am arguing against also presupposes such a distinction. The notions of analytical truth and conceptual analysis are surely connected; conceptual analysis is practice which when successful arrives at analytical truths and analytical truths are truths which are in principle reachable by means of conceptual analysis alone. Thus if no distinction could be made between analytical and non-analytical statements, then surely no distinction could be made between conceptual analysis and other activities such as explanation either, such as the contrast they make presupposes. Thus their claim will fail to be interestingly true in any case; either there is no such dividing line between analysis and explanation as they think or the problem of universals falls on a different side of the line than they think.}

Explanation Swoyer or Rodriguez-Pereyra talk about, as the targets of explanation they offer are or at least at first sight seem to be existential generalizations. However, it is clearly not the analysis of individual events (or singular cases) they are talking about, and singular cases and laws are the only kinds of explananda usually given for scientific explanation, so this just strengthens the evidence for the claim that the explanation they talk about cannot be a kind of ordinary scientific explanation or even at all similar to it.

\footnote{Nagel’s example is also commonly used to support quite dubious ontological theses, which we must distinguish from his correct thesis concerning explanation. Many philosophers would conclude from the example that dispositions like dormitive virtues cannot be used to explain anything genuinely. They would even infer from this that there is no need to postulate them in our ontology at all, though this inference is fallacious since even if dispositions would not be needed to explain anything they might yet be needed to describe something correctly. However, I do not think the example gives any kind of support even to the claim that dispositions could not be used to explain anything genuinely. We must make clear that the same fact can be a genuine explanans for one explanandum but only a pseudo-explanans for another explanandum. The existence of a disposition like a dormi-}
Even if we have doubts whether Nagel’s criterion is a good general criterion of a satisfactory explanation, it is quite compelling when it is applied to explanations that are supposed to serve as the basis of an abductive argument. The main feature which is supposed to distinguish abductive as well as inductive arguments from deductive i.e. demonstrative ones is just that in an abductive or inductive argument the premises can be true while the conclusion is false, while in a deductive argument they cannot. As the premise of an ostensibly abductive argument is the explanandum, then if the explanandum implied the conclusion analytically, the premise could not be true without the conclusion being so, and thus the argument would after all be a deductive one. Thus it is vitally important that an explanation that is to serve as the basis of an abductive argument should satisfy Nagel’s criterion for a satisfactory explanation. However, I will argue that none of the usual proposed solutions to the problem of universals do satisfy Nagel’s criterion for a satisfactory explanation and thus they are not after all reached abductively but instead deductively.

There are strong reasons to fear that many realistic solutions to the problem of universals do not satisfy Nagel’s criterion. For example, Gary Rosenkrantz says in [Ros93, page 72] that

A standard argument in favor of realism posits the existence of a property in order to explain why several particulars are of the same kind.

However, if this were truly the problem, then there are reasons to suspect that the explanans supposed to be offered by universalists such as Rosenkrantz, that there are universals such as kinds and properties, seems

tive virtue cannot indeed be used to explain genuinely (even together with other factors) why opium causes sleep, since such an explanandum already implies that opium possesses such a disposition. This is probably because the concepts of disposition and causation are not independent, but there is a connection between them, though it is controversial whether this is because causation is to be reduced to dispositions or conversely. However, the existence of a disposition like a dormitive virtue can be a part of a genuine explanation of why someone falls asleep, since this explanandum does not imply anything about dispositions. Suppose that someone who does not know anything about the effects of opium or milk sees a person who has taken opium mixed with milk fall asleep and wonders why this happened. It can be informative to explain to him that this was because opium possesses a dormitive virtue, since he can infer thence that the person fell asleep at least in part because of the opium, and exclude the possibility that it was because of the milk alone. Thus we may not infer from Nagel’s example anything concerning the ontological status of dispositions, but we can infer from it much concerning which kinds of explanations are genuine and which are not.
to be identical with the explanandum, being just the same thing expressed more technically with the coined technical term "universal". Kinds are after all - depending on how you use the word - most naturally interpreted as either the same thing as universals or some subclass of them such as Porphyry’s species and genera are. Even if the word is interpreted in the later way Rosenkrantz produces just a classification, not a genuine explanation. This would be like trying to explain the existence of electrons with the aid of the existence of material objects, which would be surely circular. Thus the Universalist’s explanation seems quite analogous to explaining why opium includes sleep by saying it has a dormitive virtue. The argument offered by Rosenkrantz could be made genuinely abductive by supposing that the kind is a second-class property in Armstrong’s sense and the property is a first-class property, i.e. a universal in Armstrong’s narrow sense. However, if what Rosankrantz says is interpreted in this way, then the opponents of realism would not accept the premise that several particulars are of the same kind, as they would not agree even that any second-class properties exist, so the argument would in any case be ineffective or even question-begging against them. We would need first a different argument to establish that several particulars share the same second-class property, and this argument could not be abductive.

Perhaps then it is no wonder that Rodriquez-Pereyra comes to a non-realistic solution; the only way to explain something like the explanandum he proposes non-trivially is to explain it away. However, Rodriguez-Pereyra’s Resemblance Nominalism does not satisfy the criteria of a satisfactory scientific explanation either.

First of all, we have a powerful argument that the Resemblance Nominalist’s explanans for a particular statement is also entailed by its explanandum. Rodriguez-Pereyra’s explanandum is that several things share a property, and as we saw he understands this so that the property in question is a

---

66 We may doubt whether a realist theory can satisfy Nagel’s desideratum even if the explanandum is supposed to be how general terms can apply to different individuals. General terms are most naturally understood as types, and types are universals, so here also the explanandum seems to imply the existence of universals, which was supposed to be the explanans. Perhaps if the explanandum could be reformulated in terms of tokens of terms alone this problem could be escaped. However, it is already a substantive issue whether such a reformulation is possible, and solving this problem may already require solving the whole problem of universals, so that nothing would remain to be explained if this could be done.
sparse property. Rodriguez-Pereyra borrows his concept of a sparse property from David Lewis and sparse properties are introduced and so implicitly defined by Lewis in [Lew86, page 59] as properties whose possession makes for qualitative similarity. Since "qualitative similarity" as Lewis uses it here is surely a synonym for "resemblance" as Rodriguez-Pereyra uses it, it follows from the very definition of sparse properties by mere deductive inference that if several things share a sparse property, then those things resemble all things with that property, which was supposed to be the explanans. It follows from that definition just as it follows from the definition of dormitive virtue that substance with such a virtue cause people to fall asleep, so Nagel would surely regard the explanation Rodriguez-Pereyra offers as a pseudo-explanation.

We can show this more technically, as this may make it more certain for those capable of following an argument conducted in symbolic logic. It follows from Lewis’s and Rodriguez-Pereyra’s implicit definition of sparse properties that $(\forall P)(Sparse(P) \rightarrow (\forall x)(\forall y)(P(x) \& P(y) \rightarrow Resemble(x,y)))$, and thus the truth of this implication belongs to the data Rodriguez-Pereyra starts out with. Therefore the data is (in the simplest case where there are two things sharing the sparse property) of the form $(\exists P)((\exists x)(x = a) \& (\exists y)(y = b) \& P(a) \& P(b) \& Sparse(P)) \& (\forall P)(Sparse(P) \rightarrow (\forall x)(\forall y)(P(x) \& P(y) \rightarrow Resemble(x,y)))$. Thus his inference has the form $(\exists P)(\exists x)(\exists y)(P(x) \& Sparse(P)) \& (\forall P)(Sparse(P) \rightarrow (\forall x)(\forall y)(P(x) \& P(y) \rightarrow R(x,y))) \vdash (\exists x)(\exists y)(R(x,y))$. This inference is surely deductively valid, and thus no abduction is needed to pass from the premise to the conclusion. Thus while the Resemblance Nominalist’s explanation may not be like explaining why opium induces sleep by saying it has a dormitive virtue, it is like explaining conversely why opium has a dormitive virtue by saying that it induces sleep, which is not any better. Here the explanans does not assert any more than the explanandum either, but only makes clearer the way we use words.

Even worse, Rodriguez-Pereyra’s ontological theory as a whole seems at least at first glance to actually contradict the explanandum. Since the explanandum resemblance nominalism offers is not logically incompatible with those universalism and trope theories are supposed to offer, Rodriguez-Pereyra must make and does make the additional claim that there are no universals or tropes so that his theory might be a genuine alternative to.
universalism. However, if the explanadum is that many things share the same property and the theory of resemblance nominalism, says there are no properties, as there are neither universals or tropes, then the explanadum and the theory explaining it surely contradict each other.

Rodriguez-Pereyra, however, sees that this counterargument might be raised and tries to answer it. Let us see if he succeeds in this. He says in [RP02, pages 15-16]:

I use the word ”property” in expressions like ”the rose has the property of being red ” or ”different particulars share the same properties” without committing myself to the existence of any entities over and above roses and particulars in general.

However, it is not evident whether Rodriguez-Pereyra can use the word like this if he uses it at all in the way it is commonly used in English. In order to see whether he can use the word like this while remaining true to the semantic rules of English he would have to give a conceptual analysis of his proposed explanandum, something he refuses to do. It can be argued that it follows from the very fact that languages like English are learnable and hence have a recursive semantics that he cannot do so without using not only the word ”property” but also the words ”the” and ”same” differently than they are usually used. However, if Rodriguez-Pereyra uses the words in an unusual way and another philosopher like Rosenkrantz uses them as they are usually used in describing his explanandum, then though the two philosophers use the same words in describing their explanandum, they really have different explananda. In this case if their problems are explanatory problems then since they have a different explanans they are different problems and the philosophers are just talking past each other. However, Rodriguez-Pereyra and Rosenkrantz still disagree with each other, since they offer incompatible solutions to the original problem formulated by Porphyry.

In fact the position of at least some Universalists, including Rosenrantz, is actually more sustainable than that of Rodriguez-Pereyra. First of all, though neither Rodriguez-Pereyra and Rosenkrantz reach their theory by reasoning abductively from their data as they claim, yet Rosenkrantz’s theory seems far better since at least there is no reason to think that it would contradict the data it starts from. Secondly Rosenkrantz tells us more about what kind of explanation he is offering than Rodriguez-Pereyra does, and it
turns out that Rosenkrantz's concept of explanation is clearly far broader than the usual concept of scientific explanation. Rosenkrantz says in [Ros93, pages 73,74] that in an explanation of the sort he is offering the explanans provides a philosophical analysis of the explanandum and that in an analysis the analysans and the analysandum are necessarily equivalent and the analysans provides a certain kind of explanation of the analysandum. If then in this kind of explanation the explanandum is equivalent with the explanans then the explanans certainly follows from the explanandum, so Rosenkrantz almost explicitly acknowledges that the kind of explanation he is offering is not informative. However, Rosenkrantz also says that the inference he is using is inference to the best explanation, but this seems highly misleading, for if in the sort of explanation he uses the explanandum and explanans are logically equivalent, then surely the explanans is the only possible explanation, not just the best of many possible explanations, as any explananda inconsistent with the explanans would also be inconsistent with the explanandum. Swoyer, however, explicitly denies that the explanation would be the only possible explanation.

The view Rosenkrantz is proposing is, then, in reality closer to the view I am here proposing than the view of Swoyer I am here opposing, though Rosenkrantz’s view is expressed a bit misleadingly. The kind of philosophical explanation Rosenkrantz seeks is explanation only in a completely different sense of the word ”explanation” than scientific explanation in Nagel’s sense. Thus Rosenkrantz cannot really be blamed because the kind of explanation he proposes does not satisfy Nagel’s criterion for a satisfactory explanation. Swoyer, however, clearly does consider the kind of explanation he is offering to be a kind of scientific explanation, and hence the fact that it does not satisfy Nagel’s criterion clearly counts against it. It is not quite clear which of these two camps Rodriguez-Pereyra belongs to. While he admits conceptual analysis as a kind of explanation, he yet contrasts the kind of explanation

---

67 Some reconciliation between these two conceptions of ontological explanation can, however, be found if we distinguish two conceptions of possibility, logical and epistemic possibility. An explanation such as Rosenkrantz seeks has to be the only logically possible explanation of its explanandum, but it might still be that there are many epistemically possible explanations, if something that is logically impossible may yet be epistemically possible, since we do not always know what is and what is not logically possible. In any case explanation in Rosenkrantz’s sense obviously cannot be contrasted with conceptual analysis as a competing methodology as e. g. Swoyer does, so using the word ”explanation” in the way Rosenkrantz does is rather misleading as it is more commonly used as Swoyer does, so that conceptual analysis could not be a form of inference to the best explanation.
he thinks the problem of universals demands with conceptual analysis. This would seem to indicate that the kind of explanation he seeks must according to him be similar to scientific explanation, unless there is to be a third kind of basically different explanation besides scientific explanation and the kind of philosophical explanation which consists in conceptual analysis; however, Rodriguez-Pereyra does not make very clear what he thinks the relationship between scientific explanation and the kind of philosophical explanation he is trying to provide to be.

Even if Rodriguez-Pereyra thinks that the explanation in question is something quite different from ordinary scientific explanation, he may be in trouble. It can be questioned whether Rodriguez-Pereyra can really formulate the problem of universals as a problem of explaining how something is the case in Nozick’s sense convincingly. If the problem is of this kind, then the existence of universals would have what he calls apparent excluders. He suggests that this excluder is just the numerical difference among particulars. He admits that some may feel that there does not seem to be any incompatibility, of any sort, between a and b being the same in kind and being numerically different. I must confess that I belong to those that feel that there is no incompatibility, at least any that could mislead any careful thinker for a second. Rodriguez-Pereyra, however, says that there is an apparent incompatibility between being the same and being different, and distinguishing between numerical difference and qualitative difference is already a way of accounting for it. Would Rodriguez-Pereyra then say that there is also an apparent incompatibility e. g. between two men having the same father and being different men, which must be accounted for by distinguishing between let us say paternal identity and numerical identity? Really no one except a philosopher motivated by some kind of monistic agenda would be inclined to confuse identity with having the same colour or shape any more than anyone except a sophistical philosopher would be inclined to confuse identity with having the same father.

I have so far examined only two possible answers to the problem of universals, though these are from the opposite ends of the spectrum of the offered solutions and so form a rather representative sample. In order to fully support my conclusion I would have to go through them all in detail, but that is clearly impossible here. However, though I cannot here go into all the details, it seems to me that if the Problem of Universals is not only
supposed to an explanatory problem but its explanandum is thought to be what Armstrong and Rodriguez-Pereyra think it to be, then all proposed solutions to the problem of universals that are at all credible make the supposed explanandum entail the supposed explanans and are hence viciously circular pseudo-explanations by Nagel’s criterion. The only solutions that do not do this are those where the proposed explanans does not explain the explanandum at all. E.g. predicate nominalism (or at least a simplistic form of it) is clearly a solution of this kind. That many things have a negative electric charge does not entail (as the deductive-nomological model of explanation would demand) that the predicate "have a negative electric charge" applies to them, as the predicate could have been used so that it applied to things with a positive charge instead or not used at all. Nor is it true that if the predicate "have a negative electric charge" would not be applied to them they would not have a negative charge as the counterfactual theory of explanation would demand. The same holds of concept nominalism (or at least a simplistic form of it); even if there had been no sentient beings and hence no concepts in the universe, there would yet have been things that had a negative charge. Summing up, all proposed explanations to Armstrong’s explanatory problem are either pseudo-explanations or wholly non-explanatory, which suggests very strongly that the whole problem is a pseudo-problem.

We cannot conclude from this that all attempts to portray the problem of universals as an explanatory problem would lead to it being a pseudo-problem. For instance, Salmieri’s Problem of Concepts does not seem to be pseudo-problem; at least some answers to it seem to give quite genuine explanations. If the universalist explains why unitary thoughts apply to a plurality of objects by the hypothesis that the plurality of objects possess a common property and that the mind is first related to that property and then through it to the plurality, this seems to be genuine explanation. The explanandum (that unitary thoughts apply to a plurality of objects) does not seem to imply the explanans (that there is a property and the mind is in some way related to it). This conclusion must be qualified a bit; the explanandum is a bit obscure, as it is not wholly clear what it means for a thought to be unitary so this is not wholly certain; someone could claim that a thought being logically unitary implies that it is related to a property, though this might not be plausible. However, the formulation could probably be only
slightly modified so that the obscurity is eliminated but Salmieri’s intent is preserved; the explanandum can just be that a thought applies to a plurality of objects. Similarly the explanatory problem could be formulated in a more linguistic way so that the explanandum is that linguistic predicates apply to several objects; this might not be a pseudo-problem either. Nevertheless, the historical argumentation I have already given in Section 6.3 suffices to show that these explanatory problems are not the same as the traditional Problem of Universals, though they are closely related problems.

6.5 A Problem of Truthmakers or of Ontological Commitments?

6.5.1 The A question and B question

Following Campbell [Cam90, §2.2, page 29], Rodriguez-Pereyra distinguishes two questions concerning universals which in his view may not have parallel answers. We can take one single red object and ask of it: what is it about this thing in virtue of which it red? Campbell and Rodriguez-Pereyra call this the A question. Secondly we can ask about two red things what is it about these two things in virtue of which they both are red. They call it the B question (note that this distinction has nothing to do with the famous distinction between A-theory and B-theory of time). Thus the A question concerns statements of the form “a is F” while the B question concerns statements of the form “a is F and b is F”. Following Oliver [Oli96, page 49], Rodriguez-Pereyra divides these questions still further into six questions which in their view Armstrong fails to distinguish. We can ask for an account of statements of the following forms:

**Example 12**  a and b have a common property

**Example 13**  a and b are both F

**Example 14**  a and b have a common property, F

**Example 15**  a has a property

**Example 16**  a is F

**Example 17**  a has the property F
Oliver and Rodriguez-Pereyra notice that Armstrong took statements of the form "a is F and b is F" and "a and b have a common property, F" to be equivalent, like statements of the forms "a is F" and "a has the property F", as have many other philosophers. It is surely intuitive that at least some equivalences of this kind are correct, though as we will see desperate nominalists have even resorted to denying this.

Oliver and Rodriguez-Pereyra point out that all of these questions can ask either for the truthmaker, the ontological commitments or the conceptual analysis of the content of the statements in question. Rodriguez-Pereyra then argues that if the Problem of Universals is the problem of giving an account of how the facts expressed by the statements are possible, then the sort of account cannot be about the ontological commitment of the statements. This is clearly true, as an impossible statement might yet have ontological commitments. However, they think that the account can be about their truthmakers, since if we show that a thing has a truthmaker we show that it is true and therefore it must also be possible.

I have some serious reservations about this conclusion. First of all, I have some doubts concerning truthmaker theory as such. Do statements (which word I use here as a proxy for whatever entities are truthbearers) in general have truthbearers? Applying the notion of truthbearers in ontology demands that some principle such as the truthmaker axiom, according to which for all statements (of some kind) there is an entity that makes them true, is correct. I have already argued in Subsection 4.2.4 that there are good reasons to doubt whether the truthmaker axiom, at least in its standard form, is true at all. I have argued that philosophical problems concerning truth can be solved on the basis of Tarski's theory of truth (if it is suitably generalized) and that this theory does not require (even in a generalized form) the postulation of any truthmakers.

However, even if the truthmaker axiom were correct, there are reasons to doubt whether the problem of universals concerns the truthmakers of some statements, so my opposition to the claim of Rodriguez-Pereyra does not depend solely on my theory of truth. Showing that a statement has a truthmaker does indeed show that it is possible (though this is surely not the only way to show that it is possible). It might perhaps even show why it is possible. However, it far from clear that it shows how it is possible. Could we not know that something is possible without knowing how it is
possible? However, this just stresses that it is not at all clear what it would be to explain how something is possible, in which case it cannot be ruled out that finding out the ontological commitments of a truth might be also involved (among other things) in showing how it is possible\textsuperscript{68}. However, the existence of a truthmaker does not prove even the possibility of truth other than in a circular way if the existence of the truthmaker is inferred only from the truth which is supposed to be the explanandum, as would be done in a typical abductive argument. If we are to have a non-circular proof we have to have independent evidence for the existence of a truthmaker, if we are to justify the possibility of the explanandum. However, if we have such independent evidence, we are not dealing with an inference to the best explanation at all, at least not if such inference is supposed to be a primitive form of inference!\textsuperscript{69}.

However, even if this conclusion were correct, it would not prove that the traditional problem of universals concerns truthmakers, since as we have

\textsuperscript{68}I would suggest very tentatively that showing how something is possible requires finding out the truth-conditions (both necessary and sufficient) of the statement whose possible truth we are asking for and then showing that they can be fulfilled. If truthmakers play any role in this, it is because the existence-conditions of a truthmaker are sufficient truth-conditions of the statement, so that the possible existence of any possible truth-maker of the statement implies the possible truth of the statement. However, if we are trying to explain just how something is possible and not how it is actual, then possible truthmakers are just as good as the actual truthmakers; we need not know what is actual in order to know what is possible. Giving the truth-conditions of the statement is in a (perhaps rather weak) sense a conceptual analysis of the statement. It also involves finding out the ontological commitments of the statement, since the existence of the entities the statement is ontologically committed to is among its necessary conditions and thus belongs to its truth-conditions.

\textsuperscript{69}It is interesting that though Rodriguez-Pereyra refers to Nozick as support for his view, Nozick has not drawn the same consequences from his view as Rodriguez-Pereyra does. Nozick tries famously (in [Noz81, pages 172-176], in the same book Rodriguez-Pereyra refers to) to explain how knowledge is possible. However, he does not do this by giving the truthmakers of statements concerning knowledge, as we should except if Rodriguez-Pereyra’s treatment of universals were correct and the problem of universals were really of the same form as the question how knowledge is possible. Rather, Nozick gives to knowledge four conditions which are supposed to be both necessary and jointly sufficient and then tries to show that these conditions can in some cases be satisfied. Thus part of what he does in order to explain how knowledge is possible can be considered as giving a paraphrase for statements concerning knowledge, or even giving a conceptual analysis of knowledge. This differs from Rodriguez-Pereyra who thinks a paraphrase need not be given for statements concerning properties in order to explain how the having of common properties is possible. Nozick’s conditions have been very much discussed in the epistemological literature, and the presupposition has always been that if the conditions can be shown not to be adequate paraphrases then Nozick’s theory of knowledge has been shown to wrong. Thus Nozick does not really give much support to Rodriguez-Pereyra.
seen the traditional problem of universals is not a problem of giving an account of how some fact is possible. Rather the traditional problem is (in its basic part) whether there are universals, and thus if we have some fact which serves as our evidence it is relevant to the traditional problem only it implies or makes probable the existence of universals. In the first case, if the evidence implies the existence of universals then it (or its verbal or propositional formulation) carries ontological commitment to universals. Even in the latter case the concept of ontological commitment comes into play, since then the evidence must make probable a statement which carries ontological commitment to universals. We come then to the conclusion(already famously defended by Peter van Inwagen in \[vI04]\) that the sort of account required is primarily about the ontological commitments of the statements rather than about their truthmakers\(^{70}\).

The account can, however, also be about the conceptual analysis of the statements. I have argued earlier that conceptual analysis i. e. meaning analysis should find out the truth-conditions of statements (instead of verification conditions or indication relations), since meaning either consists in truth-conditions or determines them. I have also argued the ontological commitments of a statement are among its truth-conditions. It may then

\(^{70}\)Oliver argues in \[Oli96, \S21\] that if Armstrong is interpreted so that he asks for the ontological commitments of the statements in question, his a posteriori realism leads him to trouble. If the argument Armstrong uses is interpreted so that it is based on the ontological commitments of the sentences in question, then a similar argument would lead to the existence of such universals as the relation of instantiation and virtues. One would have to deduce from the truth of the sentence "John has all the virtues of his father." (when understood as presupposing that John’s father has virtues) that there are virtues, However, Armstrong thinks that an a posteriori realist cannot establish the existence of virtues as easily as this. This seems to me to be a perfectly valid argument ad hominem against Armstrong. However, the problems do not really arise because Armstrong is an a posteriori realist but just because Armstrong is a sparse realist. A promiscuous realist avoids the problems, as he can easily accept the existence of virtues (at least if he is a realist about ethics). Even a promiscuous realist can be an a posteriori realist and think that we know the existence of universals only by means of experience. In order to be an a posteriori realist he just has to think that it is only by experience that we know that any such sentence as "John has all the virtues of his father" is true. This indeed seems to me to be a plausible view. Surely an a posteriori realist does not have to deny that we can deduce one statement with ontological commitments from another statement with stronger ontological commitments! Surely it is enough for him that the ultimate basis of our knowledge of what exists is empirical. Whether even a promiscuous realist can accept the existence of the instantiation relation is another matter, as this poses the danger of Bradley’s regress and (what is far more serious) even of Russell’s paradox. However, I will argue below that while some instances of quantification over universals are indispensable for our inferential practice, others can be avoided, if we make use of free predicate logic.
(at least in some cases) be by means of conceptual analysis that we find out the ontological commitments of the statement.

In fact the problem of universals would not be at all an ontological problem if it were a problem about truthmakers. According to Rodriguez-Pereyra in [RP02, page 18], all solutions to the problem of universals make claims about what exists, and that is why it is an ontological problem. However, surely making claims about what exists is not enough to make a theory an ontological theory nor the problem to which it is a solution an ontological problem. Many theories that are not ontological make claims about what exists. A physical theory may claim that there are electrons, or a biological theory may claim there are cats, but this does not turn physics or biology into ontology, though it does make physics and biology relevant to ontology. An ontological theory should concern only the most general kinds of questions with regard to what exists. Rodriguez-Pereyra says in [RP02, page 223] that Resemblance Nominalism and Class Nominalism have the same ontologies. However, if they have the same ontologies, and yet are different theories, then surely the problem to which they offer different answers cannot be an ontological problem. The problem of what are the truthmakers of some truths is not an ontological but a semantical problem. It does have implications for ontology and cannot thus be sharply separated from it, but this is just because statements about the existence of truthmakers have existential commitments just like statements of physics and biology.

Rodriguez-Pereyra then argues further that according to standard truthmaker theory whatever entities make statements of the form $16 \text{true}$ make also statements of the form $13 \text{true}$ and whatever entities make such statements true make also statements of the form $12 \text{true}$, since whatever makes a logically stronger statement true makes also are logical weaker statements, i.e., all statements implied by it true. Therefore according to Rodriguez-Pereyra, thus the different kinds of nominalism are not always different ontological theories, as is usually thought, but many of them (such as predicate nominalism and concept nominalism) are just different ways of justifying a single ontological theory (at least if the concept nominalist admits that there are (token) predicates though he does not employ them in explaining universals away, as a concept nominalist surely usually does). However, clearly some differences between different kinds of nominalism, such as the difference between a class nominalism which holds that there are classes though not intensional properties and an extreme or thoroughgoing nominalism which holds that there are neither classes nor properties, are properly ontological.
Pereyra the fundamental problem is to give the truthmakers of a statement of the form 16, e. g. what is it about a single thing that makes it red. Thus the fundamental problem is according to him the A question, not the B question. Thus in his view the One Over Many Problem dissolves into the Many over One problem, the problem of how the same thing can be of different types.

Again I agree that the argument is valid but deny one of its premises, against which I have already argued. Since as I have argued it is the ontological commitments of the statements which are relevant to the problem of universals, not their truthmakers, we must ask how the ontological commitments of the two kinds of statements are related and not how their truthmakers are related. As Oliver already pointed out, the relation of the ontological commitments of the two kinds of statements to each other is the converse of the relation of their truthmakers. The ontological commitments of the weaker statements are also the ontological commitments of the stronger statements. Whatever entities the statement "a and b have a common property" is committed to the statement "a and b are both F" is committed to if the first statement truly follows from the second and similarly whatever entities the statement "a has a property" is committed to the statement "a is F is" is also committed to. Thus we must primarily try to find out what are the ontological commitments of statements of the form 12 and thus the fundamental problem is the B question.

At first glance statements of the form 12 seem to be straightforward existential quantifications over properties that are universals, and they would hence by Quine’s and Church’s criteria of ontological commitment carry ontological commitment to universals. However, nominalists have argued

---

72Even this statement is somewhat controversial. Ross P. Cameron has argued in [Cam08] that the ontological commitments of a theory are just those things that must exist to make true the sentences of the theory. If Cameron were correct, then asking for the ontological commitments and asking for the truthmakers of the six statements would be the same thing. Ontological commitments and truthmakers can only be contrasted in the way Rodriguez-Pereyra and Oliver contrast them if we assume some theory of ontological commitments like Quine’s. Cameron proposed his theory as part of a solution to van Inwagen’s Special Composition Question, an ontological question separate from the Problem of Universals. However, it can also be applied to the Problem of Universals. If so applied it leads to pretty much the same conclusion as Rodriguez-Pereyra’s metaontological theory, that the A question is fundamental. However, there are reasons not to accept Cameron’s view: Jonathan Schaffer has presented many of them in [Sch08b]. Any formulation of the truthmaker principle makes use of quantification, so why should quantification have ontological commitments only as a part of a truthmaker principle, and never elsewhere?
that the syntactic form of such sentences systematically misleads us so that they are not really existentially quantified sentences, so we must study such claims.

Of course while this is what we must ultimately do, this may not be the best point to start from. After all, the statement "a and b have a common property." seems to imply quite trivially the existence of universals, so we might be accused of begging the question if we just assumed it to be true. Can one indeed argue at all rationally that the statement is true but there are not universals? This seems hard, as it seems obvious that the statement quantifies over properties, which are common to two individuals and hence are universals. However, according to Quine’s criterion of ontological commitment, as expressed in [Qui53d], we are ontologically committed to entities we quantify over. Alonzo Church suggested in [Chu58] the modification to this definition that we are only committed to what we quantify over existentially. However, universals are quantified over existentially in the statement, so it seems clear that the statement is ontologically committed to universals. One (very implausible) way for the nominalist to escape this conclusion would be to argue that the phrase "have a common property" is here used like "kick the bucket", where saying that "John kicked the bucket." does not really quantify over buckets as as its syntactic composition would lead us to except and hence does not carry ontological commitment to buckets. I will later show that this kind of interpretation is not at all plausible.

6.5.2 Higher-order and First-order Notation in Theories of Predication

A major reason, perhaps the only reason, we have to think that statements of the form "a and b have a common property" are true is that we think they are implied by statements of the form "a and b are both F" and that these are again implied by statements of the form "a is F" and "b is F". It would surely be absurd to claim that no statement of the latter form would ever be true, as this would imply that no atomic statement that could express empirical knowledge about the world would ever be true. Thus the only thing that a nominalist can with any show of reason doubt is whether the implication holds. As this can be questioned, we must ask whether the equivalence between "a and b are both F" and "a and b have a common
property, F” really holds as it surely seems to, as some philosophers have
denied it. This equivalence seems perfectly analogous to the equivalence
between “a is both F and G” and “the particular a is F and is G”.
Symbolically the first equivalence is between $F(a) \land F(b)$ and $(\exists P)(P = F \land P(a) \land P(b))$ if we use higher-order notation. Using such notation
is necessary if we are to use the predication operation of predicate logic to
express the same as predication expresses in natural languages, since natural
languages are inescapably higher-order in their syntax, having for example
such expressions as adverbs, which turn a predicate (an adjective or verb) to
another predicate and are thus second-order functors which can be defined
with the aid of second-order predicates which are predicated of first-order
predicates.

However, as the interpretation of any syntactic operations defined with
the aid of formation rules is conventional, it is not necessary to use the
same syntactic operation with the same interpretation in all languages, so
we can just as well introduce into the language of first-order predicate logic a
special predicate to express the predication operation of natural languages.
George Bealer has shown in [Bea82, page 84] how a formula of a higher
language can always be translated into a formula of a first order language.
This first order language is of course richer than the standard system usu-
ally known as the language of first order logic, but it is first-order in the
exact syntactic sense that no variable occurs in predicate position. Bealer’s
logic, P1, has a special predicate representing the instantiation or predica-
tion relation. Bealer uses a special predicate, $\Delta$, standing for the relation
of predication. $\langle t_1, \ldots, t_n \rangle \Delta x$ means that the relation x is predicated of the
entities $t_1, \ldots, t_n$. This predicate differs from the membership predicate as
it does not fulfill the axiom of extensionality. Bealer suggests in [Bea82, page
258] that if we adapted anadic logic such as was developed in [Gra76] where
predicates need not have a fixed adicity this could be written more simply as$t_1 \ldots t_n \Delta x$. I suggest that we can also write this in usual prefix notation as
$\Delta(t_1, \ldots, t_n, x)$ or $\Delta(\langle t_1, \ldots, t_n \rangle, x)$. Bealer also uses an abstraction opera-
tor; $[A]_{v_1, \ldots, v_n}$ means the relation that holds between objects $v_1, \ldots, v_n$ just
in case $A$ if $n > 1$, the property of being such that $A$ if $n = 1$ and the propo-
sition that $A$ if $n = 0$. Bealer calls properties, propositions and relations
together PRPs, and I will occasionally use this convenient terminology.

Bealer follows the tradition of Aristotelianism and of Lesniewski in that
this relation comprehends both instantiation and identity. Therefore \( a \Delta b \) means that \( b \) is said of \( a \) and covers also the case where \( a \) and \( b \) are identical. I will, however, depart from Bealer in that I will distinguish these two kinds of predication from each other from the beginning, as is more usually done. I will use \( \Delta \) to stand only for the instantiation relation, not for identity. Thus \( \langle t_1, \ldots, t_n \rangle \Delta x \) means that the relation \( x \) holds between the entities \( t_1, \ldots, t_n \).

However, Bealer’s translation method can be used despite this change. I will show as an example how this is done in the case of second order logic.

Thus we can also express the first ostensible equivalence between natural language sentences using first-order notation as an equivalence between \( \Delta(a, f) \& \Delta(b, f) \) and \( (\exists x)(x = f \& \Delta(a, x) \& \Delta(b, x)) \), when \( \Delta(x, s) \) means that \( s \) is predicated of \( x \), i.e. \( x \) that is instantiated by \( s \). Similarly symbolically the second equivalence is between \( F(a) \& G(a) \) and \( (\exists x)(x = a \& F(x) \& G(x)) \) or between \( \Delta(a, f) \& \Delta(a, g) \) and \( (\exists x)(x = a \& \Delta(x, f) \& \Delta(x, g)) \). We must ask whether there are any reasons to doubt this analogy.

In fact there might be such reasons. One reason to doubt the first equivalence as it is expressed in higher-order notation might be doubt about higher-order logic. However, all doubts one might have about higher-order logic are not relevant here in discussing the problem of universals. One might doubt whether the equivalence between \( F(a) \& F(b) \) and \( (\exists P)(P = F \& P(a) \& P(b)) \), i.e. \( F(a) \& F(b) \equiv (\exists P)(P = F \& P(a) \& P(b)) \), is a logical truth. The question of what truths are logical and what not is not an easy one, and has been much discussed without much agreement being reached (for a good treatment of the topic see [Mac00]).

While I find it quite plausible that the equivalence is a logical truth, this difficult question is fortunately not relevant here. It is enough in order to argue for the existence of universals that this equivalence and the material implication \( F(a) \& F(b) \to (\exists P)(P = F \& P(a) \& P(b)) \) which it implies by the definition of equivalence are truths, whether logical or non-logical truths. E.g. one could take them to be truths of a property theory or set theory that is non-logical but is an indispensable foundation of mathematics and so strongly justified, and this would be enough to support realism about universals. After all if the material implication is true at all then we can infer \( (\exists P)(P = F \& P(a) \& P(b)) \) from the material implication and \( F(a) \& F(b) \) with the aid of modus ponens alone, and modus ponens belongs already to
propositional logic, not to any kind of predicate logic, whether higher-order or first-order. In order to be at all relevant here doubts about higher-order logic must be radical, so that one doubts whether the material implication above is a truth at all; this requires some very radical position, such as doubt whether any statement containing higher-order quantification can be true or indeed whether higher-order quantification is at all meaningful.

However, if the equivalence and implication can also be expressed in first-order notation, as $\Delta(a, f) \land \Delta(b, f) \equiv (\exists x)(x = f \land \Delta(a, x) \land \Delta(b, x))$ and $\Delta(a, f) \land \Delta(b, f) \rightarrow (\exists x)(x = f \land \Delta(a, x) \land \Delta(b, x))$, as I think, then even this kind of doubt about higher-order logic is not important for the problem of universals. This material implication is surely true in standard first-order logic.

However, there may be a deeper cause for doubt. Though both of the equivalences above always hold in standard predicate logic, they do not hold universally in free logic (at least not in positive free logic), since in it either of the terms $a$ and $f$ may be non-denoting, but hold only for some formulas in free logic. Such equivalences are according to free logic true not in virtue of the form common to all of them but in virtue of their content or of some form which is the form of only some of them. Thus a philosopher who thinks that free logic is the correct predicate logic could argue that the implication holds only in the case of the inference to the existence of particulars but not in the case of the inference to the existence of universals. However, we must see if he can consistently hold this while admitting that the kinds of inference that he would use in the course of his ordinary life are valid, for if he cannot then he cannot consistently deny the inference to the existence of universals to be valid.

Thus we must consider whether any such equivalences are correct. Quine has argued that the only way to show that statements with some ontological commitments are true is pragmatically, by showing that such statements are in some way useful. This may mean that they are useful in science or (here I may depart from Quine) in ordinary life. This also holds of equivalences that imply sentences with such commitments together with other, already accepted sentences. However, an equivalence can be useful only because it allows us to infer a useful sentence from another useful sentence. Thus our task comes down to considering whether any equivalences of the form we have considered above play an important role in our inferential practices.
This is of course a big task which I cannot carry very far here; however, in the rest of this dissertation I will give some initial arguments for the conclusion that many of them do. If any such equivalence holds between an indispensable or very useful sentence of the form 13 and a sentence of the form 12 which quantifies explicitly over universals, then we must give a positive answer to the most fundamental question concerning universals, whether there are any. In the last section of this dissertation I will give a preliminary argument for the conclusion that some such equivalences hold and therefore there are universals.
A Preliminary Argument for the Existence of Universals

Many sentences of natural languages such as Example 18 below are true. What they express is the case, and is often known to be case. Indeed, if our epistemology is directly realistic, i.e., allows us to have direct knowledge about material particulars, then it is certainly also plausible that we can directly see that two shirts are of the same colour.

Example 18 These shirts are the same colour.

Example 19 The first shirt is blue.

Example 20 Blue is a colour.

Example 21 The second shirt is blue.

The first sentence is obviously similar to the following.

Example 22 John and Mike love the same woman.

This kind of sentence is usually formalized as follows:

Formalization 1 \((\exists x)(W(x) \& L(j, x) \& L(m, x) \& (\forall y)(Wy \& L(j, y) \rightarrow y = x) \& (\forall y)(Wy \& L(m, y) \rightarrow y = x))\).

It is a natural assumption that the sentence about the shirts must be formalized analogously, as quantifying over colours. This assumption is so
natural that surely anyone who suggests a different kind of formalization - such as taking the "having the same colour" to be a simple primitive expression - has the burden of proof. He must say why we cannot similarly take "loving the same woman" to be primitive, so that Example 22 would not involve commitment to women either.

This assumption can also explain why the statement expressed by the fourth sentence about the shirts follows from those expressed by the three earlier sentences. If "having the same colour" were a primitive then the correctness of the inference would be wholly mysterious. Here the assumption of the existence of universals will play an explanatory role, though this is only explanation in the general sense in which a mathematical axiom can explain something. Also the suggestion that "having the same colour" is a primitive would make the inference non-logical, but the inference plainly appears to be logical just like the sentence about love in whatever sense natural language inferences can be logically valid.

Nor would such a suggestion of primitiveness easily if at all fit in with any kind of compositional or even minimally recursive theory of semantics. It would amount to treating the phrase "have the same colour" as similar to phrasal expressions like "kick the bucket". Not only is this clearly unintuitive (and staggeringly uneconomical), but it would lead to insuperable problems with the learnability of languages. Clearly there are infinitely many phrases that exhibit similar inference patterns; e. g. "have the same pale colour", "have the same small size", "have the same many angled outline", "have the same size bigger than 3 cubic meters", etc. A language in which all such phrases were primitive would have infinitely many lexical primitives and could not be learnable.

If the sentence is to be formalized as quantifying over colours there are two ways to do this, the higher order approach and the first order approach.

The following formula is its formalization in higher order notation. In the formula below $C(F)$ means intuitively that $F$ is a colour.

Formalization 2 $((\exists F) (F(s_1) \land F(s_2) \land C(F) \land (\forall G)(C(G) \land G(s_1) \rightarrow G = F)) \land (\forall G)(C(G) \land G(s_2) \rightarrow G = F)))$.

$^1$Thus expressions like "have the same colour" are productive in a sense in which phrasal expressions like "kick the bucket" are not; an expression like "kick the yellow bucket of Farmer Ted" does not have a conventional use in which it would indicate anything concerning dying, but has only a reading in which it quantifies over buckets.
The following is the first order formalization of the same sentence. In the
formula below \( C(x) \) means intuitively that \( x \) is a colour and \( \Delta(s, x) \) means
that \( x \) is instantiated by \( s \).

\[
\text{Formalization 3 } \exists x (C(x) \land \Delta(s_1, x) \land \Delta(s_2, x) \land (\forall y)(C(y) \land \Delta(s_1, y) \rightarrow y = x) \land (\forall y)(C(y) \land \Delta(s_2, y) \rightarrow y = x)).
\]

In these formulas, however, we are quantifying both existentially and
universally over colours. Colours are properties. Therefore we are quantifying
existentially over properties. According to the formulas the properties
we are quantifying over are instantiated by (or predicated of in the ontological
sense of predication) two distinct particulars. However, properties that
are instantiated by or predicated two distinct particulars are by definition
universals according to the traditional definition we have already quoted.
Therefore we are quantifying existentially over universals in the sense I (and
many other philosophers) use the word.

Here a dogmatic Quinean might have an objection. I am here making use
of the notion of what entities are universals by the very definition of the word
"universal". Quine was of course famously suspicious about this notion of a
definition. He asked in [Qui53e, page 21] "Who defined it thus, and when?"
This may be often be a difficult question to answer, though this is hardly
sufficient evidence to think that it has not been defined by some persons,
whether explicitly or implicitly. However, in this case it can be given a
clear answer. As I have already shown, Aristotle originally defined it in this
way. I have also shown that other, newer philosophers such as Wolterstorff
and Loux have followed them in defining universals in essentially this way,
sometimes with minor differences, and could show the same of many more.
On could of course quibble about whether "universal" is a correct translation
of "katholou" or "universale", but it is a historical fact that it has commonly
been translated in this way and this has helped to constitute the meaning
of the word "universal" as it is is used in present ontological discussion (and
the English word has actually obviously developed even phonologically out
of the Latin word). How the words are used in ordinary Greek, Latin and
English or by the practitioners of other scholarly disciplines (such as e. g.
linguistics which has a rather different use for the word) is wholly beside
the point, especially since the Greek form of the word word katholou was
first coined by philosophers, probably by Aristotle himself, and originated
within philosophy, and the corresponding words in later languages derived from this. Likewise one might ask whether De Interpretatione from which the notion ultimately derives was really written by Aristotle or some of his disciples, but again this small uncertainty certainly does not matter. In fact a philosopher could just define the word "universal" stipulatively so that an entity is an universal if it is or could be instantiated by many entities, and even Quine admits that there is no problem with meaning entailments generated by such stipulations. However, if every philosopher just did this then it would not be known whether different philosophers who spoke of universals were dealing with the same question. Therefore reference to the history of the notion of a universal is useful though not necessary.

However, according to Quine’s and Church’s criteria of ontological commitment we are ontologically committed to entities we quantify over existentially, if the quantification in question is objectual. The kind of ontological commitment to universals in question here follows from the example sentences by the definition of universals and is therefore what I have earlier called narrow analytical ontological commitment (and defined in 9 and in 10).

Is the quantification here objectual? I do not have space here to discuss this question in sufficient detail, which is a large part of the reason why this argument for the existence of universals is only preliminary. However, we can note that the quantification in these examples does not appear to be any of the most common kinds of non-objectual quantification. As I have already argued, we do not have names for (all shades of) colours, so the quantification here cannot be understood as substitutional quantification. Nor do the sentences constitute any intensional context, which suggests that the quantification is not perspectival or any of the other kinds of non-objectual quantification I have discussed in Section 5.10 which may occur in intensional contexts without implying corresponding objectually quantified statements. Therefore it seems to me that if a nominalist wants to escape this argument by claiming that the quantification used in such examples is not objectual, the burden of proof falls upon him: he must tell us what kind of quantification is in question.

Therefore if the quantification in such sentences as Example 18 is objectual, as it seems to be, we are committed to the existence of universals if we ever use such sentences and hold them to be true. However, it is surely
absurd to deny the truth of such sentences. Moreover, using such sentences is necessary if we are to use such forms of inference as the inference from Example 18 and Example 19 to Example 21 above. Making use of such inferences, however, is surely essential to our practical life, so denying them is unacceptable for pragmatisable reasons. It would be rather awkward, for example, if your wife or mother asked you to buy her three shirts of the same colour and you had to answer her that you cannot do that, as you are a nominalist and do not believe in colours! Thus I conclude that we have at least very good prima facie reasons to believe in the existence of universals.

It must be noted that though the example was given using shirts, and so we presupposed the existence of shirts, it can also be given in a way that even philosophers that absurdly deny the existence of shirts cannot evade. We could start from a sentence like ”These sense-data are the same colour.” Thus the argument should be inescapable even to phenomenalists. I have argued for (weak and feeble) foundationalism, according to which there must be observation statements, and the examples I have given confirm that universals must be quantified over in some such statements, no matter whether such statements concern physical entities or sense-data or mental entities.

We could of course also start from such statements as ”These elementary particles are the same size.”. Thus the argument should be inescapable even to the kind of extreme physicalists who think that only elementary particles exist. However, such examples would in my view be less epistemologically secure, so they would make the existence of universals less certain.

It must also be noted that though the example given above proved the existence of properties in the narrow sense, entities like colours, a perfectly similar argument (though one whose premises are epistemologically less se-

---

2 Someone might object that it is only the universally quantified part of the sentences that is useful, and that the existentially quantified part is redundant. However, if we admit that the transition from the universally quantified sentence ”Whatever colour this shirt has that shirt has.” to the unquantified sentence ”If this shirt is blue then that shirt is blue.” preserves truth, we must also admit that the transition from the unquantified sentence ”This shirt is blue.” to the existentially quantified one ”This shirt has a property.” also preserves truth. As we have already seen, if an ostensible singular term can be bound by an universal quantifier it can also be bound by an existential quantifier and must thus be a genuine singular term. Surely if a transition from (\forall x)(\phi \rightarrow \psi) to \phi(a/x) \rightarrow \psi(a/x) preserves truth then a transition from \phi(a/x) to (\exists x)(x = a \& \phi) also preserves truth. This holds even in free logic, since there the correctness of the first inference already presupposes the truth of the sentence (\exists x)(x = a)!
cure) can be constructed also for other universals, which not all would be willing to call properties (though many modern philosophers call them also properties, and I mostly follow this usage in this dissertation), namely natural kinds or species or substance-universals; just consider the following sentences and the inference from the first two to the third:

**Example 23** *This animal and that animal are of the same species.*

**Example 24** *This animal is a tiger.*

**Example 25** *That animal is a tiger.*

Similar proofs can of course be given for the existence of relations. E. g. we can say that an object a is as far from object b as object c is from d, and here apparently the same spatial relation is said to hold between a and b as between c and d. Such statements seem to be needful if we are to develop any kind of scientific geometry and thus if we are to develop any kind of physics. Thus quantification over relations is as important as quantification over properties.

Similar proofs can also be given for the existence of mathematical objects such as numbers. This can be seen from examples such as the following:

**Example 26** *As many men live in Chicago as in New York.*

**Example 27** *Nine hundred men live in Chicago.*

**Example 28** *Nine hundred men live in New York.*

We should again explain how the third example sentence follows from the two preceding ones. The most obvious formalizations of these sentences that help to explain this are the following higher-order ones (where $NN(F)$ means that $F$ is a natural number):

**Formalization 4** $(\exists F)(NN(F) \& F(\lambda x)(M(x) \& L(x,c)) \& F(\lambda x)(M(x) \& L(x,n))) \& (\forall F)(NN(F) \rightarrow (F(\lambda x)(M(x) \& L(x,c)))) \equiv F(\lambda x)(M(x) \& L(x,n))).$

**Formalization 5** $9((\lambda x)(M(x) \& L(x,c))).$

**Formalization 6** $9((\lambda x)(M(x) \& L(x,n))).$
These higher-order formalizations can again be replaced with slightly more complex first-order ones:

**Formalization 7** \((\exists x)(NN(x) \& [M(y) \& L(y, c)]_y \Delta x \& [M(y) \& L(y, n)]_y \Delta x) \&
(\forall x)(NN(x) \rightarrow ([M(y) \& L(y, c)]_y \Delta x \equiv [M(y) \& L(y, n)]_y \Delta x)).\)

**Formalization 8** \([M(y) \& L(y, c)]_y \Delta 9\).

**Formalization 9** \([M(y) \& L(y, n)]_y \Delta 9\).

This example seems to me a very good example of the simplest possible way mathematical concepts are used in everyday discourse. It is surely plausible that the way they are used in sciences like physics is a generalization of this. A crucial problem in the philosophy of mathematics is how mathematics is applicable to physical objects. There are many senses in which applicability here has been understood. We can distinguish a semantic problem of applicability and an epistemic problem of applicability. The semantic problem is what are the truth-conditions of statements that combine talk about numbers with talk about non-mathematical objects such as physical objects. The epistemic problem is how we know which such statements are true. I have here suggested how the semantic problem of the applicability of mathematics is to be solved, though of course this will not solve the epistemic problem. The example above gives a first indication of how a theory of universals can explain such application in the semantic sense. As we cannot know all the men living in New York personally, we can only refer to them with the aid of such properties of theirs as 'being a man living in New York'. These are typically clearly relational properties, but such properties do not differ essentially from ostensibly intrinsic properties such as shapes or colours. Mathematical entities are primarily related in discourse to such properties of physical objects and only indirectly to their instances. The relation between mathematical entities and properties of physical objects is just instantiation, the very same relation that holds between physical objects and such first-order universals as shapes or colours.

The epistemological problems concerning mathematical objects like numbers and pure sets, such as have famously been pointed out by Paul Benacerraf e.g. in [Ben64], must be admitted to be more difficult than those concerning other universals. These problems are especially hard for an empiricist. I have argued at length that our knowledge of sensible qualities like
colours or shapes is straightforwardly empirical; we perceive them with our senses, just as we do sensible particulars. Physical properties like charges or masses on the other hand can just like unobservable particulars such as electrons be known inductively or abductively. These examples should then be unproblematic even for an atomistic empiricist. However, it is not quite clear whether mathematical objects can be known in either of these ways. Nevertheless, it is not clear that they cannot be known in any of these ways, either.

I have argued that there is prima facie evidence in favour of the existence of universals. This does not of course prove conclusively that there are universals. I ought then to argue that such evidence is not defeated; e. g. I should show that the attempts of nominalists to show that the quantification used in such examples as I have used is not objectual do not succeed.

---

3 One way to argue that they can be is of course the theory of Quine that statements concerning such objects are justified holistically. However, there may also be an atomistic explanation of how we have knowledge of mathematical objects. Penelope Maddy argued in [Mad80] that it is possible to acquire perceptual beliefs about such mathematical objects as sets, in other words to perceive them. Maddy thought that sets are not universals but individuals or particulars - a view I have argued against in Section 6.3.3 of this work - but views about how we acquire knowledge about sets are separable from views about their ontological position. That we perceive sets may at first sight seem implausible. However, it is not wholly clear what it would take for us to perceive sets. Let us remember that (as I argued in Section 6.4.1 of this work) the claim that we perceive some entities can by semantic ascent be transformed into the equivalent claim that these entities are mentioned (in such a way that we can quantify over them) in some observation statements or their existence is implied by such statements. Logic of perception (as presented e. g. in [Nii82]) seems to confirm this. According to it we perceive that there are some entities if and only if it is true in all worlds consistent with our perceptions that there are such entities. It seems extremely plausible that quantifiers ranging over mathematical entities can and must be used in describing our experience. E. g. we can say that a person sees that there are more than ten trees in front of him (or more than ten sense-data in his visual field), but he cannot say exactly how many. In order to symbolize this ascription of perception it must be understood so that the person sees that there is such a number \( m \) which is bigger than ten that there are \( m \) trees, in symbols (where \( S_a p \) means that \( a \) sees that \( p \), \( S_a ((\exists m)(N(m) \land m > 10 \land m((\lambda x)(Tx))) \) or \( S_a ((\exists m)(mΔn \land m > 10 \land [Tx]_xΔm)) \)).

This implies \( S_a ((\exists x)(N(x))) \) or \( S_a ((\exists x)(xΔn)) \), i. e. that \( a \) sees that there is a number. Maddy’s examples of perceived sets are all impure sets; she uses such examples as sets of eggs. The numbers that can be used in the description of what we experience are of course finite numbers. However, if we can gain knowledge of some mathematical entities by observation, then (as Maddy also remarks) the prospects of extending that knowledge to other mathematical entities such as pure sets or transfinite numbers by induction or abduction do not seem to be too dark. However, if this strategy does not work, I do not see that we have any alternative but to hold with the earlier logicians that the existence of mathematical objects is an analytic truth (and hence that existential truths can be analytic). Just rejecting the existence of mathematical objects is not an option because of their indispensability to our theories.
Unfortunately, there is not enough space left in his work to do this, so it must be the subject of a future work. If my argument so far has been correct, then I think I have already shifted the burden of proof to nominalists who want to deny the existence of universals; they should find out if there are any defects in the kind of analyses I have given or the argument based on them, as well as whether there exists evidence that would undermine the evidence I have presented. Since the main aim of my work was to find out what methods to use in exploring ontological problems, and especially the problem of universals, this aim is already fulfilled; I have shown what kind of evidence ought to be given for realism concerning universals, and it must be left a further question, whether this kind of evidence is sufficient and how much of it there is and how good it is.
Bibliography


508


510


[Ger] Lloyd P. Gerson. Platonism and the invention of the problem of universals. available online at http://individual.utoronto.ca/lpgerson/Platonism_And_The_Invention_Of_The_Problem_Of_Universals.pdf.


Ernst Mach. *Contributions to the Analysis of the Sensations*. Open Court, 1897. translated by S. M. Williams.


[Mor85] Donald Morrison. $\chi\omega\rho\sigma\tau\omicron\sigma$ in aristotle. *Harvard Studies in Philology*, 89:89–105, 1985.


Heinrich Scholz. Metaphysik als strenge Wissenschaft. Staufen-Verlag, Köln, 1941.


[Smi] Barry Smith. Ontology as a branch of philosophy. available online at http://mediastream.buffalo.edu/Content/research/phismith/OntologyVideoFiles/OntologyLectures09B_byLecture/Ontology09Lecture1.wmv.


[WheIX] William Whewell. *Of Induction with Especial Reference to Mr. J. Stuart Mill’s system of Logic*. John W. Parker, MDCC-CXLIX.


Index

A question, 484
abduction, 402
Abelard, Peter, 424
absolute concepts, 93, 425
absolute generality, 290
abstract entities, 435
abstraction, 57
Achinstein, Peter, 138, 403
acquisition argument, 154, 235
Adam of Wodeham, 197
aetiology, 21, 38
Aho, Tuomo, 333
Ajdukiewicz, Kazimierz, 178
Alexander of Aphrodisias, 412
Alexander’s dictum, 275
All-in-One Principle, 292
Allair, Edwin B., 126, 456
anadic logic, 259, 491
analogy, 281
analysis, 12
Appiah, Anthony, 155
Aquinas, Thomas, 178
Argument from Illusion, 126, 134
Aristotle, 20, 189, 416, 421
Armstrong, D. M., 198, 275, 409, 433, 459
Arnauld, Antoine, 10
Asay, Jamin, 241
Austin, J. L., 62
Austin, J. L., 250
Avenarius, Richard, 69, 106
Avicenna, 22, 412
Ayer, Alfred Jules, 96, 115, 220, 285
Azzouni, Jody, 308, 341, 378, 417
B question, 484
Bacon, John, 393, 464
Baker, G.P., 154
Balaguer, Mark, 230, 276, 277
Barnes, Jonathan, 423
baroque ontology, 402
Bealer, George, 491
behaviourism, 119, 120
being as being, 20
Belnap, Nuel, 187
Benacerraf, Paul, 275, 501
Bergmann, Gustav, 39, 109, 126, 273, 435, 455
Bergmann, Julius, 179
Berkeley, George, 97, 105
Berlin Circle, 89
Berlin, Isaiah, 115
Berti, Enrico, 282
Beth, Evert W., 10
Black, Max, 215
Bocheński, Józef M., 30, 35
Boethius, Anicius Manlius Severinus, 412, 434
Bogdanov, Alexander Aleksandrovich, 107
Bolzano, Bernhard, 218
Bonjour, Laurence, 88, 136
Boolos, George, 293, 400
Bradley, Francis Herbert, 97, 273
Brandom, Robert, 366
Bressan, Aldo, 332
Bronstein, Eugene, 46
Burgess, Alexis, 225
Burley, Walter, 424
Cameron, Ross P., 489
Campbell, Keith, 39, 471
Candlish, Stewart, 205
Carnap, Rudolf, 89, 285, 317, 425
Cartwright, Richard, 290
Cassirer, Ernst, 76
categories, 20, 45
causal theory of denotation, 218
causal theory of knowledge, 277
causal theory of perception, 79, 84, 134, 462
Chakrabarti, Kisor, 421, 457
Chakravartty, Anjan, 49, 138, 145
Chalmers, David J., 161
Chateaubriand, Osvaldo, 387
Chihara, Charles S., 298, 385
Chisholm, Roderick, 191
Chomsky, Noam, 133, 150
Chomsky, Noam, 154
Church, Alonzo, 115, 352
Classical Model of Science, 10, 36
classification, 45
Clatterbaugh, Kenneth C., 126
Clauberg, Johannes, 26
Cocchiarella, Nino B., 35, 66, 69
Coffa, Alberto J., 88
Cohen, S. Marc, 29
cohesion theory, 211
Colyvan, Mark, 277, 279
composition, 12
compositionality, 383
computer science, 161
Comte, Auguste, 69
conceptual analysis, 485
conceptual roles, 230
conceptualism, 417
concrete entities, 435
conservativeness, 256
Convention C, 356
Convention T, 353
conventionalism, 71, 135
problems, 88
Corazzon, Raul, 26
core-dependent homonymy, 281
correspondence rules, 137
correspondence theory of truth
fact-based and object-based, 182
weak and strong, 248
cosmology, 20
Cracow Circle, 30
Creath, Richard, 115
Crivelli, Paolo, 189
Crusius, Christian August, 31
Damnajovic, Nic, 205
David, Marian, 182, 192, 195
Davidson, Donald, 11, 173, 224, 225, 378, 459
de Jong, Willem, 10, 36
deflationary theory of truth, 168
deflationism
about meaning and content, 222
about metaphysics
mild, 161
strong, 161
about truth, 222
methodological, 227
mild, 34, 175
demonstration, 36
Demopoulos, William, 79
Descartes, Rene, 7
descriptional attributes, 439
descriptive metaphysics, 119
Dever, Josh, 383
Devitt, Michael, 8, 11, 51, 121, 169, 185
Dewey, John, 306
direct realism, 102, 465
Disquotation Schema, 355
Disquotational Theory of Truth, 355
Dolson, C. Daniel, 5
Duhem, Pierre, 17, 98, 137
Dummett, Michael, 120, 151, 225
Effertz, Dirk, 58
Einstein, Albert, 78
Eklund, Matti, 287
Eleatic Principle, 275, 303
eliminativism (about truth), 216
Ellis, Brian, 145, 428
emergentism, 121
empirical realism, 67
epistemological priority, 8
Etchemendy, John, 65, 392
Eudoxus of Cnidus, 13, 414
Evans, Gareth, 460
Ewing, A. C., 113
exemplification, 426
existential monism, 281
explanation, 409
Fales, Evan, 277, 458
fallibilism, 87
Fechner, Gustav Theodor, 306
Feigl, Herbert, 120, 134, 154
Fetzer, James, 128
fictional objects, 350
fictionalism, 417
Field, Hartry, 215
Fine, Gail, 436
Fodor, Jerry, 139
formal mode of speech, 90, 472
formalism, 89
foundationalism, 11, 74
feeble, 136
weak and strong, 87
foundherentism, 136
frameworks, 320
free predicate logic, 318
Frege, Gottlob, 51
French, Steven, 77
Freuler, Leo, 26
Friedman, Michael, 71
Frost-Arnold, Greg, 217
functionalism, 121, 228
fundamentalism, 40, 443
Follesdal, Dagfinn, 333
Gödel, Kurt, 91, 457
Gardiner, Mark Quentin, 155
genera, 420
general metaphysics, 26
Gerson, Lloyd P., 411, 414, 423
Gibson, Roger, 140
Gilson, Etienne, 30
Glanzberg, Michael, 179
Goclenius, Rudolf, 26
Goldman, Alvin, 231, 278
Goodman, Nelson, 405, 439, 447
Gottlieb, Dale, 309
Gower, Barry, 76
Gregory of Rimini, 197
Grice, H.P., 141
Grossmann, Reinhardt, 457, 469
Grover, Dorothy, 187, 310
gru, 439
Guendling, Jack, 5
Gupta, Anil, 332
Guendling, Jack, 5
Gupta, Anil, 332
Hacking, Susan, 11, 52, 306, 327
Hacker, P. M. S., 154
Haddock, Guillermo, 62
Hale, Bob, 66, 161
Hanson, Norwood Russell, 404
Hansonian abduction, 404
Harman, Gilbert H., 404
Harmanian abduction, 404
Haukioja, Jussi, 154
Heck, Richard G., 66
Hegelianism, 97
Heidegger, Martin, 8, 28, 95
Heil, John, 8
Hempel, Carl G., 128, 137, 473
Henry of Harclay, 418
hermeneutics, 313
Hilbert, David, 89
Hintikka, Jaakko, 332, 378, 426, 460
Hirsch, Eli, 281, 287
Hochberg, Herbert, 126
Hodges, Michael P., 341
Horwich, Paul, 187, 193
Husserl, Edmund, 37, 179, 435, 457
Ibn Sīnā, 22, 412
ideal being, 285
identity conditions, 282
identity criteria, 297
Identity Theory (of mind and brain), 120, 219
incommensurability, 137
Indian philosophy, 421, 457
indication relations, 230
individual domains, 392
individual essences, 430
individuals, 448
induction, 439
infinitism, 40
information
semantic and epistemic, 160
Ingarden, Roman, 37, 281
instantiation, 426
instrumentalism, 147, 307
internal and external questions, 320
internal realism, 67
interpretation functions, 392
intersubjectivity, 132, 134
intuition, 101
intuitionism, 151
Jaeger, Werner Wilhelm, 23
James, William, 304, 464
Jones, J. R, 425
Jubien, Michael, 282, 297
Kölbel, Max, 225
Künne, Wolfgang, 182
Kahn, Charles C., 426
Kant, Immanuel, 7, 22, 37
Katz, Jerrold, 51
Katz, Jerrold J., 62
<table>
<thead>
<tr>
<th>Name</th>
<th>Page Numbers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Keinänen, Markku</td>
<td>433</td>
</tr>
<tr>
<td>Kemeny, John G.</td>
<td>391</td>
</tr>
<tr>
<td>Ketland, Jeffrey</td>
<td>256</td>
</tr>
<tr>
<td>Kivinen, S. Albert</td>
<td>434</td>
</tr>
<tr>
<td>knowledge by description</td>
<td>125</td>
</tr>
<tr>
<td>Knuuttila, Simo</td>
<td>426</td>
</tr>
<tr>
<td>Kolář, P.</td>
<td>208</td>
</tr>
<tr>
<td>Koskinen, Heikki J.</td>
<td>304</td>
</tr>
<tr>
<td>Kotarbiński, Tadeusz</td>
<td>179, 218</td>
</tr>
<tr>
<td>Kretzmann, Norman</td>
<td>197</td>
</tr>
<tr>
<td>Kriegel, Uriah</td>
<td>40</td>
</tr>
<tr>
<td>Kripke, Saul</td>
<td>55, 125, 218, 310, 364</td>
</tr>
<tr>
<td>Kuhn, Thomas S.</td>
<td>138, 144</td>
</tr>
<tr>
<td>Ladyman, James</td>
<td>53, 77</td>
</tr>
<tr>
<td>Lambert, Karel</td>
<td>318</td>
</tr>
<tr>
<td>Lance, Mark</td>
<td>225</td>
</tr>
<tr>
<td>Laudan, Larry</td>
<td>78</td>
</tr>
<tr>
<td>Lenin, V. I.</td>
<td>107</td>
</tr>
<tr>
<td>Lepore, Ernest</td>
<td>139</td>
</tr>
<tr>
<td>Lewis, David</td>
<td>436</td>
</tr>
<tr>
<td>Lewis, Frank</td>
<td>423</td>
</tr>
<tr>
<td>Lewis, Peter J.</td>
<td>79</td>
</tr>
<tr>
<td>Linnebo, Øystein</td>
<td>290</td>
</tr>
<tr>
<td>Livingston, Paul</td>
<td>61</td>
</tr>
<tr>
<td>locational attributes</td>
<td>439</td>
</tr>
<tr>
<td>logical atomism</td>
<td>42</td>
</tr>
<tr>
<td>logical empiricism</td>
<td>69</td>
</tr>
<tr>
<td>logical positivism</td>
<td>69</td>
</tr>
<tr>
<td>logicism</td>
<td>59</td>
</tr>
<tr>
<td>Lorhard, Jakob</td>
<td>26</td>
</tr>
<tr>
<td>Lotze, Hermann</td>
<td>179</td>
</tr>
<tr>
<td>Lowe, Jonathan E.</td>
<td>428</td>
</tr>
<tr>
<td>Lutz, Sebastian</td>
<td>110, 115</td>
</tr>
<tr>
<td>Lvov-Warsaw School</td>
<td>218</td>
</tr>
<tr>
<td>Münster Group</td>
<td>31</td>
</tr>
<tr>
<td>MacFarlane, John</td>
<td>65</td>
</tr>
<tr>
<td>Mach, Ernst</td>
<td>69, 106, 230</td>
</tr>
<tr>
<td>Maddy, Penelope</td>
<td>276, 469, 502</td>
</tr>
<tr>
<td>manifestation requirement</td>
<td>154, 235</td>
</tr>
<tr>
<td>Manley, David</td>
<td>161</td>
</tr>
<tr>
<td>Marcus, Ruth Barcan</td>
<td>309</td>
</tr>
<tr>
<td>Markosian, Ned</td>
<td>39</td>
</tr>
<tr>
<td>Martin, C. B.</td>
<td>8</td>
</tr>
<tr>
<td>Marxism</td>
<td>21, 107</td>
</tr>
<tr>
<td>mass terms</td>
<td>354</td>
</tr>
<tr>
<td>material mode of speech</td>
<td>90</td>
</tr>
<tr>
<td>Matthews, Gareth B.</td>
<td>281</td>
</tr>
<tr>
<td>Maxwell, Grover</td>
<td>76, 81, 85</td>
</tr>
<tr>
<td>McGee, Alan</td>
<td>290</td>
</tr>
<tr>
<td>meaning postulates</td>
<td>60, 391</td>
</tr>
<tr>
<td>Meinong, Alexius</td>
<td>194, 285, 345</td>
</tr>
<tr>
<td>Merlan, Philip</td>
<td>30, 33</td>
</tr>
<tr>
<td>metaethics</td>
<td>5</td>
</tr>
<tr>
<td>metalanguage</td>
<td>311</td>
</tr>
<tr>
<td>metametaphysics</td>
<td>5, 6</td>
</tr>
<tr>
<td>metaontology</td>
<td>5</td>
</tr>
<tr>
<td>metaphysical priority</td>
<td>8, 109</td>
</tr>
<tr>
<td>metaphysics</td>
<td></td>
</tr>
<tr>
<td>Aristotel’s definition</td>
<td>20</td>
</tr>
<tr>
<td>Ayer’s definition</td>
<td>97</td>
</tr>
<tr>
<td>Bradley’s definition of</td>
<td>97</td>
</tr>
<tr>
<td>immanent</td>
<td>49, 107</td>
</tr>
<tr>
<td>Michael, Michaelis</td>
<td>301</td>
</tr>
<tr>
<td>Milkov, Nikolai</td>
<td>179</td>
</tr>
<tr>
<td>Mill, John Stuart</td>
<td>52, 403</td>
</tr>
<tr>
<td>Misak, Cheryl</td>
<td>124, 466</td>
</tr>
<tr>
<td>model theory</td>
<td>210, 392</td>
</tr>
<tr>
<td>Moderate Realism</td>
<td>412</td>
</tr>
<tr>
<td>Modest Account of Truth</td>
<td>189</td>
</tr>
<tr>
<td>Montague, Richard</td>
<td>378</td>
</tr>
<tr>
<td>Montague, William Pepperell</td>
<td>463</td>
</tr>
</tbody>
</table>
Moore, George Edward, 20, 102, 158, 179, 273
Mora, Jose Ferrater, 26
Moreland, J. P., 435
Morrison, Donald, 21
Mulligan, Kevin, 201, 212, 464
Myers, William T., 306
Nagel, Ernst, 475
naive theory of ontological commitment, 358
Natorp, Paul, 23
naturalism, 52, 121
  aposteriorist and scientific, 52
  reformist and revolutionary, 52
Neo-Aristotelianism, 8, 428
Neo-Fregeanism, 66, 161
Neo-Kantianism, 23, 71, 76
Neo-Scholasticism, 13, 30
neo-verificationism, 277
Neoplatonism, 28, 411
Neurath, Otto, 92, 107, 132, 248
neutral monism, 106, 218, 285
Newman, M. H. A., 79
Newton, Isaac, 78, 403
NF, 292
Niiniluoto, Ilkka, 79, 173, 206, 332, 460
noemata, 347
nominalism, 417
  abundant property nominalism, 437
  class nominalism, 450
  predicate nominalism, 418
Norton, Bryan G., 71, 320, 329
  nuclear vs. non-nuclear properties, 349
Nyāya-Vaiśeṣika philosophy, 421, 457
object language, 311
observation sentences, 130
Ockham, William of, 418, 419
Oddie, Graham, 79, 275
Old Academy, 414
Oliver, Alex, 484
One over Many, 416
ontic communism, 434
ontological assay, 39
ontological commitment, 485
  explicit, 298
  implicit, 298
  implicit formal, 374
  narrow analytical, 390, 498
  pragmatic, 299, 301
  relative, 392
  syntactic, 308
ontological pluralism, 287
ontological relativity, 330
ontological supercommitment, 381
ontology
  definition, 25
ontophobia, 159
ontosophy, 26
ordinary language philosophy, 157
ousiology, 21
Owen, G. E. L., 425
Owens, Joseph, 30
Paavola, Sami, 404
Parmenides, 340
Parsons, Josh, 199, 212
Parsons, Terence, 348
particularism, 448
Passmore, John Arthur, 70, 145

546
Patterson, D., 246
Paul, L. A., 56
Peacock, Howard, 298, 391
Peacocke, Christopher, 117, 155, 460
Peano, Giuseppe, 66, 449
Peirce, Charles Sanders, 40, 41, 304, 306
Pereira, Benito, 26
pessimistic induction, 78
Petzoldt, Joseph, 106
phenomenology, 37, 417, 457
physicalism, 217, 239
Pihlström, Sami, 49, 304, 306
Pikkert, Owen, 288
Place, U. T., 120
Place, Ullin T., 134
Plantinga, Alvin, 192
Plato, 275, 416
pneumatology, 25
Poincare, Henri, 71
Pols, Edward, 102, 321
Porphyry, 411, 434
positivism, 69
possibilism, 31
pragmatic realism, 67
pragmatism, 304, 464
predicables, 432
predication, 444, 491
preunderstanding, 313
Price, Huw, 70, 305, 322
primitivism, 174, 224
Principle of Acquaintance, 125
principle of non-contradiction, 32
Principle of Tolerance, 88, 136
Prior, Arthur, 88
private language argument, 154
Problem of Concepts, 408, 483
problem of non-existence, 340
projectible, 405, 439
properties
abundant, 445
first class, 443
second class, 242, 443, 460
sparse, 445
third class, 443
propositional theory of perception, 459
propositions
as signs, 197
Fregean, 193
Russellian, 191
prosentential theory of truth, 187, 237, 310
protocol sentences, 130, 472
proxy functions, 81, 150
Psillos, Stathis, 41
Putnam, Hilary, 67, 117, 287, 296
quantification
non-objectual, 313
objectual, 311
perceptual, 332
perspectival, 332
substitutional, 308, 312
unrestricted, 290
quarks, 149
Quine, Willard van Orman, 86, 88, 110, 120, 144, 268
Raatikainen, Panu, 140, 153
radical empiricism, 464
Ramsey sentences, 81
Rayo, Agustín, 291
Reale, Giovanni, 21

547
realism
  lightweight, 161
  promiscuous, 41
  structural, 76
  ontic, 77
realism (concerning universals), 418
  class realism, 450
Platonistic and Aristotelian, 276
Received View (of scientific theories), 138
recursivity, 383
Redhead, M. I. G., 81
reduction, 215
reflection, 108
Reinach, Adolf, 179
relativism, 120, 169, 286, 296
representation theory of truth, 195
resolution, 12
Rijk, Lambertus Marie de, 283, 314
Rodriguez-Pereyra, Gonzalo, 201, 402, 409
Rorty, Richard, 306
Roscelin, 418
Rosenkrantz, Gary, 477
Ross, Don, 52, 53, 77
Routley, Richard, 344
Russell, Bertrand, 42, 72, 106, 125, 136, 179, 185, 284, 449
Ryle, Gilbert, 286
Salmieri, Gregory, 408, 414
Schaffer, Jonathan, 39, 443, 489
Scheler, Max, 61
Schlick, Moritz, 124, 250, 268
scholasticism, 412
Scholz, Heinrich, 10, 31
scientism, 52
Scotus, Duns, 282
Seigfried, Charlene Haddock, 306
semantic ascent, 90, 472
semantic closure, 383
semantic holism, 129, 137, 300
semantic locality, 383
Semantic View (of scientific Theories), 138
semirealism, 85
sensibilia, 468
Shapiro, Stewart, 258
Simons, Peter, 170, 201, 212, 464
simplicity, 383
Sirkel, Riin, 415
Smith, Barry, 8, 201, 212, 464
Soames, Scott, 173
Sorabji, Richard, 197
Spade, Paul Vincent, 425
Special Composition Question, 489
special metaphysics, 26
species, 420
Spurrett, David, 53
Stace, W. T., 118
states of affairs, 179, 190
stimulus, 133, 228
Strawson, Peter, 119, 141, 157
structuralism, 76, 149
structures, 80
abstract, 81
concrete, 81
Stubenberg, Leopold, 106
Stumpf, Carl, 179
subsistence, 284
substantivalism (concerning space-time), 230
supervaluationism, 381
Suppe, Frederick, 117
Swoyer, Chris, 402, 409, 452

Synechism, 41
Syntactic View (of scientific theories), 138
Syntacticism, 89
Synthesis, 12
Syrianus, 28

T-schema, 353
Tahko, Tuomas E., 8
Tarski, Alfred, 91, 168, 177, 291
Textor, Mark, 461
Theology, 22
Theories, 300
Thom, P., 194
Tipler, Clemens, 26
Tramel, Peter, 136
Transcendence, 97, 101
Transcendental realism, 67
Transcendentals, 282
Translation, 355
Tredennick, Hugh, 33
Trinitarianism, 419
Tropes, 425, 464
Truth conditions, 220
Truthmaker, 199
Truthmaker axiom, 199
Truthmaker essentialism, 199
Truthmaker maximalism, 201
Truthmaker necessitarianism, 199
Truthmaker principle, 199
Truthmakers, 198, 485
Truthmaking, 198
Tulenheimo, Tero, 339

Universals, 421

Definition of, 421
Univocalism, 280
Upper ontology, 161
Use theory of meaning, 154
Uzquiano, Gabriel, 291

Vallicella, William F., 273
Van de Putte, M. M., 61
Van Fraassen, Bas C., 78, 403
Van Inwagen, Peter, 6, 281, 350, 487
Verificationism, 70, 220, 456
Vienna Circle, 218

Walter Chatton, 418
Wehrle, Walter E., 411
Weir, Alan, 290
Whewell, William, 403
White, Morton, 280
Whitehead, Alfred North, 449
Wiggins, David, 297
William of Champeaux, 424
Williams, Michael, 225
Williamson, Timothy, 290, 297
Wittgenstein, Ludwig, 42, 80, 113, 179
Woleński, Jan, 248
Wollf, Christian, 21, 27
Wolterstorff, Nicholas, 431, 465
World-lines, 332
Worrall, John, 76
Wright, Crispin, 66, 161
Wulf, Maurice de, 413