A VINDICATION OF LOGICAL NECESSITY AGAINST SCEPTICISM
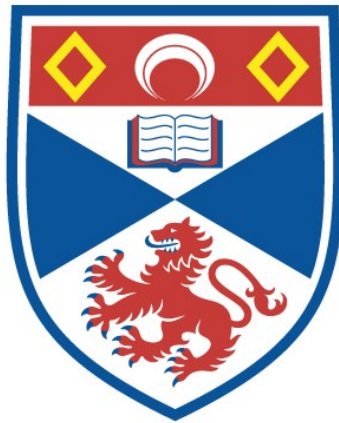
Patrice Philie

**A Thesis Submitted for the Degree of PhD
at the
University of St Andrews**

2002

THE UNIVERSITY OF ST. ANDREWS

A VINDICATION OF LOGICAL NECESSITY AGAINST SCEPTICISM

A DISSERTATION SUBMITTED TO
THE EXAMINING COMMITTEE IN CANDIDACY FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

DEPARTMENT OF LOGIC & METAPHYSICS

BY

PATRICE PHILIE

ST. ANDREWS, SCOTLAND

MAY 2002

ProQuest Number: 10170795

ProQuest 10170795

π
$\in$ 172

DECLARATIONS

(i) I, Patrice Philie, hereby certify that this thesis, which is approximately 73.000 words in length, has been written by me, that it is the record of work carried out by me and that it has not been submitted in any previous application for a higher degree.

date 21/05/2002 signature of candidate

(ii) I was admitted as a research student in September 1997 and as a candidate for the degree of Ph.D. in Philosophy in September 1997; the higher study for which this is a record was carried out in the University of St Andrews between 1997 and 2002.

date 21/05/2002 signature of candidate

(iii) I hereby certify that the candidate has fulfilled the conditions of the Resolution and Regulations appropriate for the degree of Ph.D. in Philosophy in the University of St Andrews and that the candidate is qualified to submit this thesis in application for that degree.

date 21/5/2002 signature of supervisor

In submitting this thesis to the University of St Andrews I understand that I am giving permission for it to be made available for use in accordance with the regulations of the University Library for the time being in force, subject to any copyright vested in the work not being affected thereby. I also understand that the title and abstract will be published, and that a copy of the work may be made and supplied to any *bona fide* library or research worker.

date 21/05/2002 signature of candidate

Dedicated to Sochetra

ABSTRACT

Some philosophers dispute the claim that there is a notion of logical necessity involved in the concept of logical consequence. They are sceptical about logical necessity. They argue that a proper characterisation of logical consequence — of what follows from what — need not and should not appeal to the notion of necessity at all. Quine is the most prominent philosopher holding such a view. In this doctoral dissertation, I argue that scepticism about logical necessity is not successful. Quine's scepticism takes three forms. Firstly, he is often interpreted as undermining, in his classic paper 'Two Dogmas of Empiricism', the very intelligibility of notions such as meaning, necessity, and analyticity. If the notion of necessity is meaningless, it is clear that ascriptions of logical necessity are also meaningless. In the thesis, I defend Quine's criticism of these notions by situating it in its historical context and emphasising that the real target in those writings is not the intelligibility of these notions as such, but only their Platonistic interpretation. I agree with Quine that a good theory about meaning, necessity, or analyticity must avoid such an ontological commitment. Secondly, Quine advocates, in the same paper, a holistic picture of knowledge and claims that in this picture, ascriptions of logical necessity are superfluous. I then show that holism à la Quine is committed to admit the necessity of statements of logical consequence. Thirdly, there is Quine's substitutional account of logical consequence (as exposed in his (1970)). He contends that this theory makes no use of logical necessity, thus showing its superfluousness. I show that any plausible account of logical consequence needs to appeal to logical necessity, thus undercutting Quine's claim — and, more generally, undercutting scepticism about logical necessity.

ACKNOWLEDGEMENTS

In the course of writing this thesis, I have been helped and encouraged by many friends, colleagues, and teachers. My supervisor, Crispin Wright, made me realise the depth and richness of the problem of logical necessity. I benefited immensely from his intensity and his intellectual integrity. Especially in this respect, he set an example that I hope to emulate in my career. I would also like to record an important debt to Stewart Shapiro, whose availability and open-mindedness were crucial for the completion of this project. The five months I spent in Ohio State University under his supervision were particularly crucial for the development of my ideas. The postgraduate community in St. Andrews has always been, for me, an invaluable source of support and encouragement. I consider myself lucky to have been a Ph.D. student in a time when so many gifted colleagues were there. The intellectual and collegial atmosphere among us proved to be a constant stimulant. My stint as a member of the 'Southgait Discussion Group' contributed to greatly improve key sections. For this, and for a multitude of other things, I wish to thank Patrick Greenough, Jesper Kallestrup and Duncan Pritchard. In the last years of my Ph.D., the postgraduates in the Hut were all helpful in many, many ways. I am thinking here of Philip Ebert, Iwao Hirose, Gerry Hough, Kent Hurtig, Cyrus Panjvani, Nicolaj Pedersen, Simon Robertson and Marcus Rossberg. I regard all of them as friends for life. Simon proof-read the thesis, and for this I will always be grateful. He accomplished this considerable task with a remarkable combination of seriousness and good humour. I also wish to take this opportunity to thank my mother, without whom the completion of this thesis would have been impossible. Finally, but in this case most importantly, my wife, Sochetra, gave me indefectible

support. My gratitude for her sacrifices, commitment, and positive attitude cannot possibly be conveyed in words.

TABLE OF CONTENTS

INTRODUCTION

0.1 Rationality and necessity

What does it mean to be rational? One mark of rationality has to do with reasoning. Reasoning, it could be said, consists mainly in our inferential practice, i.e., in correctly inferring propositions from other propositions. For instance, it is not *rational* to infer that Socrates is mortal from the premise that the Earth is round. On the other hand, it *is* rational to infer that Socrates is mortal from the premises that all men are mortal and that Socrates is a man. The central notion of this feature of rationality is that of logical consequence, a notion very closely tied with logical necessity. Indeed, it seems at first blush that instances of correct logical consequence such as:

(Premise 1) All men are mortal.

(Premise 2) Socrates is a man.

(Conclusion) Socrates is mortal.

hold of necessity. That is, it seems that if (1) and (2) are true, then the truth of (C) necessarily follows. It also seems, intuitively, that the truth of such ascriptions of logical necessity is something we recognise, something that is imposed on us — it appears that we have cognitive access to logical necessities and that we are equipped to 'track' these modal facts. These two thoughts are manifestations of, respectively, a non-sceptical stance towards logical necessity and a realist outlook.

Thus we have it that:

1. Reasoning is central to rationality.

2. The notion of logical consequence — of 'what follows from what' — is (partially) constitutive of our inferential practice.

3. Logical necessity is, at first blush, a crucial property of correct statements of

logical consequence.

It follows from this that logical necessity is an important ingredient of at least one important feature of rationality, namely correct reasoning. An enquiry into the nature of logical necessity can therefore be seen as a contribution towards an understanding of rationality. This thesis is such an enquiry, approaching the topic via a thorough study of the various kinds of scepticism brought against the notion of logical necessity.

In this introductory Chapter, we will first circumscribe the object of the present thesis — logical necessity — in a more rigourous manner than has been done with the intuitive characterisation given above. It will also be shown that the concept of logical necessity manifests itself in different ways. Following that, various kinds of scepticism about logical necessity will be introduced. Each of these sceptical stances will constitute, roughly, a chapter in the body of the dissertation. The goal of the thesis is to defend the notion of logical necessity against these sceptical arguments. In the last section of this Introduction, the plan of the thesis is outlined.

0.2 Circumscribing the topic: logical necessity

Let us now try to give a precise characterisation of logical necessity. Something necessary *tout court* is something that, to borrow Quine's expression, is true 'come what may'. The notion of necessity often appears in a qualified fashion. Philosophers talk about *epistemic* necessity, *metaphysical* necessity, *physical* necessity, *logical* necessity, and the like; moreover, they contrast these different kinds of necessity. Some writers on the subject reject some of them while retaining others, and some go as far as rejecting all of them. It is also possible to classify these kinds of necessity according to their relative strength. There is an interesting argument to the effect that logical necessity is the 'strongest' kind of necessity, that it is, so to speak, the *absolute* kind of necessity. It is impossible, without anticipating what follows in the course of this

thesis, to discuss whether such a classification is correct or even appropriate.[1]

The present investigation will concern logical necessity exclusively. The *nature* of logical necessity is a question that must be addressed right at the beginning of such an enquiry. Of course, to characterise the nature of logical necessity is not to give a philosophical account of it, but is merely to specify the object of our study. Which account such a notion of logical necessity deserves and whether logical necessity, so conceived, exists or not, are the key questions addressed in this study.

What, then, is logical necessity? First, a word on what it is not. Logical laws are often said to be necessary and, accordingly, there is a temptation to focus the issue upon them. That is to say, one might wish to conduct an investigation about logical necessity by comparing different sets of logical systems and weighing the reasons for adopting this or that set of logical laws. For instance, one might study the dispute between classicists and intuitionists, and arrive at the conclusion that, say, classical logic is the 'right' logic. On this basis, one could maintain that logical necessity is exemplified by classical logic, and that what is logically necessary is determined by it. The present inquiry shall not be concerned with this kind of dispute. In fact, the most interesting and penetrating sceptical arguments about logical necessity go beyond the examination of a particular logical system and its properties. What they are concerned with is the question: what makes a correct theorem of *any* logical system necessarily follow from its axioms?

Let us make this claim more precise. The logical necessity in play here is closely related to the concept of proof. Any axiomatic logical system consists of a set of axioms with rules of inference. The theorems of a particular logic are said to follow from these axioms by repeated use of the rules of inference. Not only that, but moreover they seem to follow *necessarily* from the axioms and the rules of inference: that is to say, the theorems of the system *must* be what they are — a logical system char-

---

[1]See Hale (1996) for a discussion on absolute and relative necessities.

acterised in the usual fashion cannot contain different theorems. They are obtained by proof, and each step in the proof follows necessarily from what has been proved antecedently. Whether these theorems are 'true' or 'correct' as such (independently of the logical system) need not concern us here — this is why the issue has nothing to do with the appropriateness of one given logical system over another. In fact, the concept of logical necessity we are interested in is not confined to its application to logical systems, but is also in play in mathematics and, more generally, in considerations having to do with 'following a rule'. In order to see this, consider that we want to say that the theorems follow necessarily from the axioms (via a suitable proof) in virtue of the *rules* of inference. The rules of the system play a crucial role. I will say more on these different manifestations of logical necessity in what follows.

Before giving a precise formulation of the concept of logical necessity, another of its general feature must be highlighted. On any account, logical necessity is a property ascribed to true statements. It is fairly uncontroversial that some statements are true; it so happens that some of them are not merely true, but are also logically necessary — unless the sceptic is right. So logical necessity is a modal operator that can be attached to sentences.[2] In the course of this thesis, therefore, logical necessity is to be understood as a sentence-forming operator, attaching to a sentence to form another sentence. For any sentence S, then, we can form the sentence $\Box$S, which reads: 'it is logically necessary that S'. That sentence in turn will be one that can be either true or false. So logical necessity is something we *ascribe* to a sentence. Whether or not we have any reasons to make such ascriptions — and, if we do, what their nature is — are the main topics to be explored in the following Chapters.

An ascription of necessity is, we have just seen, of the form '$\Box$S', where 'S' is a sentence and '$\Box$' the sentence-forming operator meaning 'it is logically necessary

---

[2]More precisely: necessity, possibility, and impossibility are *alethic* modalities, in that they affect the sense in which a sentence is true or false.

that'. Now of course, such ascriptions can be either true or false. To correctly (truly) apply the operator '□' to a sentence 'S', we need to know the range of sentences to which '□' is susceptible to engender truth. It is, for instance, perfectly meaningful to say 'it is logically necessary that the cat is on the mat', but also obviously false. In order to keep things as clear as possible, let us concentrate on the first manifestation of logical necessity, namely the one exemplified in logical systems. In this case, what we have is an axiomatic base, together with rules of inference. Given the axioms and given the rules, we want to say that the theorems (logically) follow. It could be represented thus:

(S*) Given any axiomatic basis A supplemented with a non-empty set of rules of inference $R_1, \ldots, R_n$, the theorems $T_1, \ldots, T_n$ follow.

Logical systems are clearly one area where, intuitively, it makes sense to apply the operator of logical necessity. But there are other instances of logical necessity, and it is important to isolate and capture these since they illustrate the pervasiveness of the concept. In his famous 'rule-following considerations', Wittgenstein often refers to the series '2, 4, 6, 8, ...', obtained by applying the function '+2' (starting with 0).[3] It seems, intuitively, that the terms of the series follow necessarily, that is to say, that each term of the series is determined in advance: it is impossible that the series turns out to contain other terms than those it actually contains. For instance, it is necessary that from 1000 onwards, the series extend to '1002, 1004, 1006, ...'. The kind of necessity illustrated in this example can be characterised — just like in the case of logical systems — as *logical* necessity. This is due to the fact that in both cases, the following ingredients play a crucial part: a basis, a rule, an outcome. In order to see this clearly, let us generalise the example as follows:

(S**) Given a number N to start with and a rule of addition R, the series $Z_1, \ldots, Z_n$

---

[3]See his *Philosophical Investigations*, §§143–242, and numerous passages in his *Remarks on the Foundations of Mathematics*.

follows.

The similarity with (S*) should be obvious.

The third example of logical necessity can perhaps be regarded as a generalisation of the second. Arithmetic is often said to be necessary. For instance, '7 + 5 = 12' is often considered to be a necessary statement. It is impossible that, for instance, '7 + 5' yields a different result than the one it actually does. Here again, the necessity in question is logical. The basis, in this particular case, consists of the two numbers 7 and 5. The rule is the rule of addition, and the outcome (12) is the result of performing the operation of addition upon the basis, namely 7 and 5 . It is clear that addition is but an example of the various operations that can be performed on numbers when one is doing arithmetic. Generalising, we find that what is logically necessary in arithmetic are statements of the form:

(S***) Given appropriate numbers $N_1, \ldots, N_n$ and a function F, the result Z follows.

Thus, to say that, for instance, '7 + 5 = 12' is a necessary statement, is to say: it is necessary that if we take the two numbers 7 and 5 and we perform addition on these numbers, 12 will be the outcome of the operation.

Now these three manifestations of logical necessity all share the same form. Let's focus first on logical systems. Take 'Γ' to refer to the axioms of an arbitrary logic, 'L' to the rules of this system, and 'Θ to refer to the theorems of this logic. Under these stipulations, (S*) has the following abstract form:

(AF) $(\Gamma \vdash_L \Theta)$

True statements of the form (AF) are exactly those, intuitively, that are logically necessary. It is this kind of statement, when we are concerned with logical necessity as it occurs in logical systems, to which the sentence-forming operator of necessity is attached to in the last analysis. It does not matter what 'Γ' is, nor what 'Θ' consists of. All that matters is that the theorems do indeed follow from the axioms. In other

words, the fact that from a given logical system certain theorems follow is not what is at issue: what is at issue is whether true statements of the form exemplified by (**AF**) are necessary or not. Sentences such as (**AF**) form the range of sentences that can be said to be logically necessary. It is easily seen that (**S****) and (**S*****) are also instances of (**AF**).

Notice furthermore that our original example:

(Premise 1) All men are mortal.
(Premise 2) Socrates is a man.
(Conclusion) Socrates is mortal.

is also an instance of (**AF**). Indeed, the 'theory' in this case consists of the premises and the 'theorem' is the conclusion. The 'logic' consists of the rules inherent in our inferential practice (in this case, modus ponens). The present inquiry is concerned, at a very general level, with getting clear about the nature of our inferential practice.

Ultimately, then, we shall focus on the necessity of sentences having the form (**AF**). That is to say, the issue of scepticism about logical necessity revolves around the meaningfulness and truthfulness of statements having the form:

(**M**) $\Box(\Gamma \vdash_L \Theta)$

Most philosophers agree that statements displaying the form of (**AF**) are capable of being true or false. Therefore, the bone of contention has nothing to do with the truth-aptness of statements exemplifying (**AF**), but rather with the truth, or correctness, of statements displaying the form (**M**). This is tantamount to saying that the issue has nothing to do with the correctness of a particular logical system, since 'Γ' can be *any* logic. It also becomes clear, when focusing on statements having the form (**M**), that the issue does not revolve around metaphysical necessity. This latter kind of necessity is the one displayed in statements such as 'water is necessarily H_2O'. For

one thing, such statements do not have the form displayed by (M). Furthermore, that kind of statement seems best understood, *prima facie* at least, as an expression of essentialism. In contrast, no such *prima facie* commitment to essentialism is involved in statements of the form (M).[4]

It is obvious enough that logical necessity, so conceived, is what distinguishes a proof from an experiment. A conditional 'B → T' — taken as formalising the trilogy basis, rule (operations carried), outcome — can easily be thought of as giving a picture of empirical experiment. An experiment, in effect, consists of an initial state on which some operations are performed, and these operations result in an outcome. Such an experiment is of course contingent, not necessary. In contrast, a proof, according to the traditional conception, is supposed to conclusively establish its conclusion in such a way that it is impossible for a correct proof to lead to a different result from the one it actually arrives to. In the case of an empirical experiment, the most it can achieve is support and adduce strong empirical evidence for a particular scientific theory. In other words, ascribing necessity to a conditional of the form 'B → T' — conceived as an experiment — is usually not an option. Contrast this with the readiness with which we accept that true statements of the form $(\Gamma \vdash_L \Theta)$ are logically necessary. Indeed, the three paradigms of logical necessity presented above ((S*), (S**), (S***)) can in effect all be understood in terms of proof, and this is shown by the fact that,

---

[4]Jackson (1998: 68-85) challenges the distinction here between metaphysical necessity and logical necessity. He argues that there is only one notion of necessity, and that notion is sufficient to explain both 'water is necessarily H_2O' and statements of the form $\Box(\Gamma \vdash_L \Theta)$. The apparent distinction between logical and metaphysical necessity can be explained, Jackson concedes, in a purely epistemological manner: the former is *a priori*, the latter *a posteriori*. But at bottom, both kinds of necessity are different manifestations of the same phenomenon and can be given a uniform account. However, Jackson's unified account of necessity relies directly on the acceptance of both propositions *and* possible world semantics. The first kind of scepticism studied (Radical Irrealism) rejects propositions and possible world semantics. Whether it is right to do so is another matter, but we can't prejudge the issue against the Quinean if we want to be fair. Moreover, even if Quine is wrong and that possible world semantics is intelligible, it is another (controversial) issue whether it makes sense to explain logical necessity on the basis of this semantics (this issue is addressed in the Appendix). Admittedly, it may well turn out that, after all is said and done against the sceptic, logical necessity and metaphysical necessity are after all manifestations of one kind of necessity. But such a view is certainly not the obvious one: it has to be earned.

intuitively, it makes sense to ascribe necessity to ($\Gamma \vdash_L \Theta$) as applied to them. Proof, then, seems to be an essential feature of logical necessity, in this sense: correct *proofs* are logically necessary, if anything is. In contrast, an empirical experiment cannot be said to be logically necessary. Whether or not it makes sense to ascribe necessity to proofs, and whether proofs really differ from experiments, are of course questions pressed by a sceptic about logical necessity.[5]

One who wishes to enquire further in the nature of logical necessity will study the concept of proof. It is one thing to notice, as I did in the last paragraph, the connection between necessity and proof; it is another to identify what exactly makes proof the object of ascription of logical necessity. One could say that given the basis Γ, the theorems Θ follow since they are *valid*. Validity, in turn, is explained in modal terms — that is to say, it is *impossible* that given Γ, the theorems do not follow. Impossibility being the mirror-image of necessity, the central question has to do with the nature of that kind of impossibility. One can therefore understand the present investigation as being concerned with defending the notion of validity against various kinds of scepticism about it, and with sketching an account of the notion of validity. We now introduce those various kinds of scepticism.

0.3 Introducing scepticism about logical necessity

Dummett (1959: 169) famously wrote that 'the philosophical problem of necessity is twofold: what is its source, and how do we recognise it?' Such a characterisation of the problem points to interesting questions and debates about the nature and the epistemology of logical necessity. It assumes, however, that there is such a thing as logical necessity[6] — that there are true (or at least, if one feels queasy about truth,

---

[5]For more on the relation between proof, experiment, and logical necessity, see Wright (1980: Part Three, especially Chapter XVII).

[6]As pointed out by Hale (1989: 175).

'correct') ascriptions of the form $\Box(\Gamma \vdash_L \Theta)$.[7] Of course, if it turns out that there is no such thing as logically necessary statements, Dummett's questions would be empty. Accordingly, the most fundamental problem of logical necessity is: 'is there such a thing as logical necessity?'. To answer in the negative is to be a sceptic about logical necessity. Reviewing various sceptical theories and defending the viability of the notion of logical necessity against them is the main objective of this thesis.

Scepticism about logical necessity, in its broadest sense, is the view according to which there is no such thing as logical necessity. When we make ascriptions of necessity, it seems at first blush that we are saying the kind of thing that is sometimes true, sometimes false. All the sceptical views studied here recoil from this and suggest that such ascriptions are *never* true. Thus the cluster of sceptical theories to be examined in this thesis have it that there are no true ascriptions of logical necessity — no true statements of the form $\Box(\Gamma \vdash_L \Theta)$. What distinguishes the various forms of scepticism are the explanations as to why such ascriptions are never true.

Broadly speaking, there are two forms such an explanation might take: Radical Irrealism and Error-theoretic Irrealism. The first is the most destructive and bold sceptical argument. It consists of the view that the notion of logical necessity itself is not meaningful. That is to say, according to this sceptical stance, ascriptions of necessity such as $\Box(\Gamma \vdash_L \Theta)$ simply do not make sense; when one judges a statement to be necessary, one would not be making a meaningful judgement — let alone a true one. This incredibly counter-intuitive thesis has been put forward, it is widely believed, by Quine in his 1951 paper 'Two Dogmas of Empiricism'. It is true that Quine's target in that paper is not explicitly the notion of logical necessity as it is understood here. Nevertheless, it is clear that his arguments apply equally well to it, since when Quine attacks analyticity, he intends it as an attack on necessity

---

[7]It also assumes that logical necessity is something we *recognise*, a claim disputed by noncognitivists (of which more below).

(in general) as well. So it will be straightforward to reconstruct his charge against analyticity as one against logical necessity as well.

There is nothing more destructive that can be said, for any notion, than claiming that it is not meaningful. If a sceptic maintains that a particular notion does not even satisfy certain accepted standards of meaningfulness, it will follow that all the other kinds of scepticism — for that particular notion — are validated, as it were, by default. Until the other kinds of scepticism have been reviewed, however, it is impossible to make that last claim precise. Yet, intuitively, it seems right to maintain that, at the very least, such a drastic charge against a particular notion is radical. This is why this form of scepticism will be called *Radical Irrealism*.

The second form of scepticism acknowledges, contrary to Radical Irrealism, that ascriptions of necessity are meaningful. It also recognises their truth-aptness: that is, on such a view, ascriptions of necessity *can*, in principle, be true or false. The sceptical thought, in this case, is not that the notion is not meaningful but rather that in all ascriptions of necessity, a false statement is produced. Statements of the form $\Box(\Gamma \vdash_L \Theta)$ are meaningful and truth-apt, but they are all false. Think of scepticism about wizards: for all we know, it is possible — in principle — that wizards exist, and to say of someone that he is a wizard is meaningful. It's just never true. This view is clearly sceptical about wizards. Similarly, the corresponding thought about logical necessity is sceptical. If such a view is true, that would mean that there is no such thing as true ascriptions of necessity. This kind of scepticism will be labelled *Error-theoretic Irrealism*, since it always erroneous, on this view, to ascribe necessity to a statement.

It is important to distinguish between Error-theoretic Irrealism and a related cluster of theories about logical necessity. These other theories also acknowledge the meaningfulness of ascriptions of logical necessity. Thus those taking this line can hold, contrary to the Radical Irrealist but like the Error-theorist, that necessity is a

meaningful notion — however, and crucially, for them it is not *truth-apt*, contrary to the Error-theorists. That is to say, ascriptions of necessity, on that account, are not expressions capable of being true or false. For instance, to say that it is necessary that $2 + 2 = 4$ is not to say something which is appraisable in terms of truth or falsity. This kind of theory need not — and usually does not — claim that '$2 + 2 = 4$' *per se* is not a true statement. What it denies is that ascribing the necessity operator over and above these true statements results in the kind of sentences which can be meaningfully true or false. According to this line of thought, then, when the necessity operator is attached to a truth-apt statement (of the kind exemplified by $(\Gamma \vdash_L \Theta)$), it results in a non-truth-apt statement. Such views are *non-cognitivist* theories of logical necessity.[8]

Expressivism and projectivism are examples of non-cognitivism. In a sense, they are compatible with Error-theoretic Irrealism since they also imply that no ascriptions of logical necessity are true. However, they differ in their diagnosis of why no such ascriptions are true. For non-cognitivists, this is due to a category mistake: ascriptions of necessity cannot be true since they are not the kind of statement that can meaningfully be true or false. Nonetheless, the non-cognitivist continues, it is possible to appraise ascriptions of necessity in other categories than those of truth and falsity. The Error-theorist, by contrast, locates the failure of truth-aptness of ascriptions of logical necessity elsewhere — namely that they are false. Expressivist views are not sceptical of the notion of logical necessity. Rather they recoil from a *realist* understanding of logical necessity. Such views will not be studied here — we are only interested in scepticism about logical necessity as such, not scepticism about a particular understanding of the notion. However, scepticism as understood here also rejects realism about logical necessity — it's just that it rejects, on top of that, *any* account of logical necessity, be it realist or non-cognitivist.

---

[8]Simon Blackburn holds such a non-cognitivist view in his 'Morals and Modals'.

Indeed, it is fair to say that scepticism also recoils from realism. For Quine, a post-positivist who developed the core of his views in the forties and fifties, non-cognitivism was apparently not a viable option. For him, statements are either truth-apt or meaningless, thus leaving no space for a non-cognitivist account — the latter requiring (at least for Quine) both a denial of truth-aptness for the class of statements under question *and* acceptance of their meaningfulness. Moreover, for him, if a discourse is truth-apt, then it is a realist discourse. Indeed, it requires a contemporary framework to acknowledge the possibility of a discourse being truth-apt and anti-realist (non-cognitivism, of course, being a version of anti-realism) — certainly a post-Quinean framework, such as the one suggested by Crispin Wright in his *Truth & Objectivity*.[9]

0.4 Plan of the thesis

The topic of this thesis, then, is scepticism about logical necessity as understood above. We have seen that sceptical theories come in two fashions: Radical Irrealism and Error-theoretic Irrealism. Both types of scepticism are prominent in Quine's writings, and his ideas will be the main target of the thesis. The Chapters of the thesis can be divided into three main parts: there is first Radical Irrealism, studied in Chapter 1. Then, in Chapters 2 and 3, we examine Quine's holism, which can be seen as a form of Error-theoretic Irrealism. Finally, Chapter 4 addresses another kind of Error-theory, namely Quine's substitutional account of logical consequence. The main textual source for Chapters 1–3 is Quine's 'Two Dogmas of Empiricism'. For Chapter 4, we approach Quine's scepticism via his *Philosophy of Logic* (1986). Nonetheless, although these will be the principal texts upon which I focus, interpreting Quine will sometimes requires understanding them within the wider perspective of his other works.

---

[9]In this section I borrowed the framework and terminology of that book to put scepticism about logical necessity in perspective.

The first Chapter deals with Radical Irrealism. Not many philosophers are interpreted as claiming that logical necessity is a *meaningless*, or unintelligible, notion. The most prominent figure seemingly advocating such a view is Quine. Indeed, he is often interpreted as holding that the notions of meaning, necessity, synonymy, and analyticity are meaningless. On this understanding of his views, it appears that Quine is a Radical Irrealist not only about logical necessity but also about meaning, analyticity, etc. This is consistent with the fact that many interpreters of Quine understand him as being a 'meaning-sceptic'.[10]

The basis for such an interpretation of Quine resides mainly in his highly influential paper 'Two Dogmas of Empiricism'. The dogmas in question are, respectively, the analytic / synthetic distinction and the verificationist theory of meaning. It is in the first dogma that his Radical Irrealist stance can be found. His attack on analyticity has a domino effect on other notions, in such a way that a critical stance towards analyticity also becomes, *ipso facto*, a critical stance towards meaning and necessity. The first thing to do, then, is to carefully study Quine's argument against analyticity in 'Two Dogmas'. I do this in the first part of Chapter 1, where an interpretation of the first four sections of that paper is given. In the course of doing that, it will become clear that Quine is not really a Radical Irrealist, contrary to what is often claimed. Moreover, despite the rather critical position taken towards Quine's views in the other Chapters, I find myself agreeing with his views on meaning and analyticity — assuming that my interpretation is faithful to what he actually meant.

The remainder of the Chapter examines various attempts to criticise Quine's argument. Grice & Strawson's counter-arguments are studied, followed by Wright's and Boghossian's. The main line of thought in these discussions is this: under my interpretation of Quine's views — backed up by references to other writings than 'Two Dogmas' — these counter-arguments do not address them. Indeed, it is clear

---

[10]For instance, Grice & Strawson (1956) and Boghossian (1997).

that they interpret Quine as a Radical Irrealist. But since he does not advocate that view, it follows that the criticisms levelled against him simply miss their target.

In Chapter 2, the first strand of Quine's Error-theoretic Irrealism — his holism — is then analysed. The textual basis for this analysis is the latter part of 'Two Dogmas'. The picture of knowledge advocated by Quine is sceptical towards logical necessity since it 'does without necessity', which is exactly what scepticism of this kind is all about. This Chapter begins, like the previous one, with a textual analysis of the relevant passages of 'Two Dogmas'. It will become clear that in this case, Quine *really is* advocating a sceptical theory about logical necessity (and, more generally, about intensional notions). Following that, the status of logical laws in holism — crucial to understand Quine's Error-theoretic Irrealism about logical necessity — is examined. This attempt to locate logical laws in the web of belief reveals a surprising fact about Quine's philosophy of logic: holism cannot accommodate logical laws as they are understood by Quine. The problem is that there is an inconsistency between Quine's epistemological views ('naturalized epistemology') and his philosophy of logic. The discussion on logical laws and holism is continued with a study of Lewis Carroll's regress in his 'What the Tortoise said to Achilles'. Here we go beyond Quine's philosophy and address crucial questions about the epistemology of logical laws in general — not only in relation to Quine's system. We close the Chapter with a discussion of Stewart Shapiro's interpretation of Quine's views regarding the a priori.

Chapter 3 resumes the discussion of holism with the study of a well-known argument by Wright. He contends that Quine's holism does not yield a satisfying philosophy of logic. According to Wright, Quine's holism collapses because he does not accept that statements of logical consequence are a priori and logically necessary. Wright's argument and various criticisms levelled against it are first discussed. Crucially, I examine a criticism to the effect that Wright's argument does not accomplish its goal (showing the necessity of necessity). I then offer an interpretation of Wright's

argument showing that it has the resources to answer that criticism. The upshot is that Wright's argument, thus understood, does show the necessity, as it were, of logical necessity.

Turning next to another kind of error-theoretic scepticism, Quine advocates a theory of logical truth and logical consequence that, he claimed, did not require any appeal to modal notions such as necessity and impossibility. This is difficult to conceive: how can a plausible non-modal account of logical consequence be possible? Indeed, the usual definition of validity given in logic textbooks and elsewhere is modal — an argument is valid if and only if it is *impossible* for the premises to be true and the conclusion false.[11] A sceptic about logical necessity, if he wants to keep the notion of validity, has to find a way to give an account of logical consequence free of modal notions. Quine does so with his substitutional account. In Chapter 4, the substitutional account is examined. It will be shown that Quine's theory does not work. More generally, it will be argued that *any* account of logical consequence — of what follows from what — has to be irreducibly modal.

---

[11]Or: an argument is valid if and only if the conclusion *necessarily* follows from the premises.

CHAPTER 1

RADICAL IRREALISM

1.1 Introduction

The most destructive and bold sceptical line against logical necessity is that the notion is not meaningful. I have called this thesis 'Radical Irrealism' since it is a version of irrealism — logical necessity, on this view, not being a *real* property — espousing a radical standpoint. Indeed, Radical Irrealism about logical necessity, if true, has the disastrous effect of depriving of meaning sentences of the form $\Box(\Gamma \vdash_L \Theta)$ in virtue of the thesis of the compositionality of meaning. If such sentences are sheer nonsense, it follows with certainty that logical necessity is not truth-apt, let alone an objective feature of the world. This thesis appears, at first blush, incredible. How is it possible that a notion we seem to understand relatively well, and for which there seems to be a reasonably uniform use in the community of competent speakers, turns out to lack meaning? How can a convincing argument to that effect be constructed and, at least potentially, lead the users of the notion to abandon it?

Quine, it is commonly assumed, is a proponent of Radical Irrealism about logical necessity. A student of Carnap and a son of logical positivism, Quine can be interpreted as having standards of meaningfulness even more stringent than those advocated by the members of the so-called 'Vienna Circle'. It is well known that logical positivists did not disregard intensional notions such as meaning and analyticity. Consider Carnap, for instance. His whole system, as displayed in *The Logical Syntax of Language*, rests on these notions, particularly the notion of analyticity. For Quine, acceptance of these notions is strongly connected to a pervasive 'dogma' of empiricism, the analytic / synthetic distinction. But, according to him, this distinction does not satisfy adequate standards of meaningfulness. A proper epistemology, then, must reject the distinction and construct our cognitive relation to the world without it. Quine's rejection of the distinction has far-reaching effects: it implies that a complete set of intensional notions are to be rejected along with analyticity. Quine's argument has a domino effect.

The notions in question include meaning and, crucially for us, necessity. For Quine, analyticity, meaning, necessity, and synonymy form an inter-definable circle of notions. The circle is so tight that, according to Quine (implicitly at least), it suffices to show that one member of the circle is meaningless to show that the others are as well. If analyticity is shown to be meaningless, then so is necessity.[1] His 1951 paper in which he spells out that view, 'Two Dogmas of Empiricism', has had such an impact on analytic philosophy — an impact still strongly felt today — that it is appropriate to devote a whole chapter on it in order to see what exactly the argument is and to evaluate it properly. Similarly to the Chapters that follow, the aim here is not wholly destructive, even if the view defended is, in the end, non-sceptical. Very often, the fact that counter-intuitive arguments have found such sympathy in the philosophical community is a good sign that there is something 'right' in them, or at least that they show something of significance. Accordingly, I will, in the course of showing where I think Quine's argument goes wrong, try to see where I think he's right. This way of approaching the various sceptical arguments will reveal suitable constraints that an account of logical necessity must meet to be both plausible and safe from sceptical attacks.

The strategy adopted to tackle Radical Irrealism as espoused by Quine will be the following. In the first place, I give a detailed exposition of Quine's argument in the first four parts of 'Two Dogmas', where he advances his reasons for saying that we can't make sense of analyticity. This thorough examination is necessary, firstly, to understand all the assumptions in play at the heart of the argument, and secondly, to see how his Radical Irrealism about analyticity is also a Radical Irrealism about necessity. Another reason for such a detailed account resides in the fact that Quine's argument is often referred to but is nevertheless seldom given a meticulous analysis.

[1]At the end of section 1, I will explain the link between Quine's argument against analyticity and logical necessity. A thorough examination of the argument is necessary to fully grasp the link

Secondly, I examine and discuss some well-known criticisms of Quine's argument. Grice and Strawson's reaction is dissected into two related parts and Wright's counter argument to Quine is also discussed in relation to Grice and Strawson's. A discussion of Boghossian's *reductio* argument against Quine's meaning-scepticism follows in the fifth section. It will emerge that despite the fact that I take none of these arguments to show conclusively that Quine's argument is flawed, they do reveal its limitations and, by the same token, its force. I conclude that our study of Quine's argument shows something important about the form any viable account of (logical) necessity must take.

1.2 Sections 1-4 of 'Two Dogmas'

One of Quine's explicit targets in 'Two Dogmas' is the analytic / synthetic distinction. However, that paper both contains a criticism of the verificationist theory of meaning, and a picture of knowledge that claims to do without the dogmas. The way he presents his argument, it appears that his alternative epistemology somehow follows naturally from the abandonment of the dogmas. That is, Quine sometimes write as if his epistemological holism and the rejection of the two dogmas stand or fall together. However, since it is possible in principle — and also in practice — to disentangle these two main strands, I will tackle them separately. This Chapter is concerned solely with Quine's attack on analyticity in the first four sections of the paper, that is, the sections that contain his argument to the effect that analyticity is a meaningless notion. In the following Chapter, I address the line of argument contained in the last two sections of 'Two Dogmas'.

Quine's overall strategy in sections 1–4 is to examine various attempts that have been made to explain the notion of analyticity and to reject them one after the other. The first attempt he considers is the following. An analytic statement is often said to be a statement which is true by virtue of meaning and independently of fact. For

instance, 'Kennedy was shot in Dallas in 1963' owes its truth both to what the words used in the statement mean and to empirical facts; but an analytic statement such as 'no bachelor is married' is said to owe its truth exclusively to the meaning of the words used in that statement, and not at all to the facts — except of course facts about meaning. The key notion presupposed in this explanation is the notion of meaning. 'No bachelor is married' is said to be true, on this account, by virtue of the meaning of 'bachelor', 'married', and the logical particles. Since 'bachelor' has the 'same meaning' as 'not married', it follows that 'no bachelor is married' is true solely by virtue of this fact about meaning. For Quine (1948: 11-12), this account, as it stands, requires that we reify meanings. Talk of '*having the* same meaning as', '*the* meaning of', and so forth, encourages the thought that meanings are entities of some sort. One could be tempted to assume that Quine is here endorsing the thesis of the priority of syntactic category: if an expression belongs, irreducibly, to the category of singular terms, then the sort of objects denoted by these singular terms exist.[2] That is, if a singular term appears in a true sentence, there seems to be a case to be made to the effect that the object denoted by the singular term exists. Quine, however, doesn't make much of this thesis: "we can use singular terms significantly in sentences without presupposing that there are the entities which those terms purport to name" (1948: 12). For Quine, then, one's ontology has to be decided on other grounds then syntactic category. If there happens to be an obstinate singular term that purportedly refers to an unwanted kind of objects, Quine uses Ramsey's device and manipulates the syntax of the language, thus transforming singular terms in another 'neutral' category.[3] The way to understand Quine's remarks about the reification of meanings in relation to expressions such as 'the meaning of' is simply that this kind of talk tempts one to reify meanings. Using such expressions makes it

---

[2]For more on the thesis of the priority of syntactic categories, see Wright (1983: chaps. 1 and 2).

[3]A very good example of this manoeuvre is in Quine (1960: 37-38), where he eliminates singular terms altogether.

natural to reify meanings, but, Quine warns us, it is a temptation we must resist.

In sum, if we try to explain analyticity by using the notion of meaning as a primary and irreducible concept in the explanation, we are led, intuitively at least, to reify meanings.[4] This might explain the temptation to succumb to an ontology of meanings — but we'll come back to that in later sections. In any case, Quine does not accept meanings in his ontology, and his main reason is certainly that it is very difficult to state a criterion of identity for meanings. We see here one of Quine's most famous slogan at work, 'no entity without identity'. In 'Two Dogmas', he does not argue directly against meanings understood as abstract entities, but relies instead on arguments he develops elsewhere.[5] A more thorough examination of Quine's demands for a satisfactory elucidation of meaning will be given in the next section, in which we scrutinise criticisms addressed to his arguments against analyticity.

1.2.1 Possible worlds and truth by definition

Let's now resume our exposition of 'Two dogmas'. After mentioning that meanings are obscure entities, Quine draws a distinction between two kinds of analytic statements. The first group consists in the logical truths, such as 'no unmarried man is married'. The truth of this statement does not depend on what the non-logical words mean, but owes its truth to the fact that it "remains true under all reinterpretations of its components other than the logical particles" (1951: 23). This is interesting since this account of logical truth does not seem to rely on meaning: Quine does not say that 'no unmarried man is married' is true by virtue solely of the meaning of the logical particles. In fact, he offers here a substitutional account of logical truth. For

---

[4]If on the contrary we try to explain meaning by using another notion, e.g., synonymy, then we have to give an account of the notion used in the explanation, and the problem of elucidating analyticity shifts from using meaning to using something else in the explanation. More on that below.

[5]See for instance Quine (1948: 9; 1960: chaps. 1 and 2 and 42-43; 1986 chap. 1.)

Quine the analyticity of these statements is not problematic (at least in that paper) since they are not true by virtue of meaning or any other unclear notion. Due to their unproblematic nature in the context of making sense of analyticity, Quine does not discuss logical truths in the first four sections of 'Two dogmas'. We shall follow his lead and focus on the second group of analytic statements. The analyticity of logical truths will be examined in Chapter 4.

The second group of analytic statements are those, like 'no bachelor is married', which can be transformed into a logical truth by putting synonyms for synonyms. Arguably, this understanding of analyticity results from an attempt to explain the notion of meaning by using the notion of synonymy: two words have the same meaning if and only if they are synonymous. Accordingly, we can replace 'bachelor' with 'unmarried man' since they are synonymous. So instead of saying that 'bachelor' and 'unmarried man' have the same meaning, we here say that they are synonymous, and that their relation of synonymy is precisely what gives 'no bachelor is married' its analyticity. Alas, for Quine the notion of synonymy is as unclear as the notion of analyticity. The problem with synonymy, considered as explaining analyticity, is akin to the problem encountered above concerning meaning: as such, the notion of synonymy is not clear and therefore it needs clarification. Quine decides to postpone discussions of synonymy and considers other explanations of analyticity.

One such explanation consists in using Carnap's notion of state description as developed in *Meaning and Necessity*, which is intended to explain necessary and analytic truths.[6] This idea is an adaptation of Leibniz's possible worlds and Wittgenstein's possible states of affairs.[7] For sake of simplicity, we'll use the more familiar notion of possible worlds in what follows — it does not affect the substance of the argument

---

[6]See Carnap (1956: 8–9).

[7]I refer, of course, to the Wittgenstein of the *Tractatus*.

to proceed that way.[8] On this account, 'no bachelor is married' is analytic since it is true in every possible world. This means that there is no world where someone is both a bachelor and a married man. This is plausible enough. But why is it the case that such a world does not exist? At this stage, the only answer must be that one will never encounter a man who is married and a bachelor since 'bachelor' *means* 'unmarried man'. So instead of throwing light on the concept of meaning and analyticity, we see that the appeal to possible worlds merely pushes back the problem to the notion of meaning. Admittedly, this way of putting the matter differs from Quine's own. Quine does it the other way round: he asks us to consider a possible world (*i.e.*, a state description) where the value True is assigned to both 'John is married' and 'John is a bachelor'. This would render the analytic statement 'no bachelor is married' false — but clearly this statement is not false. Viewed this way, the problem is: why can't we assign the value True to both statements about John? The friend of analyticity is forced to answer: because 'no bachelor is married' is *true by virtue of meaning*, thereby blocking the possibility of a world which would render this statement false. We thus see clearly that the explanation of analyticity in terms of possible world (or state description) presupposes the notion we are trying to explain, as is shown by the occurrence of the italicised idiom 'true by virtue of meaning' in the previous sentence.

A further attempt to explain analyticity involves claiming that 'no bachelor is married' is true by definition (section 2 of Quine's paper). The thought is that this statement is analytically true since 'bachelor' is *defined as* 'unmarried man'. But what is the basis of such definitions? Quine rightfully points out that an appeal to the dictionary won't do in this case. The lexicographer is an "empirical scientist" who bases his definitions on usage. He correlates 'bachelor' and 'unmarried man' from his

---

[8]It is assumed, for the sake of the argument, that we have a clear and unproblematic semantics for possible worlds. A sceptical about necessity as understood in terms of possible worlds would challenge this assumption.

observation that these two words are used synonymously. This empirical discovery of an anterior relation of synonymy between two expressions cannot explain, in the relevant way, why it is permissible to substitute 'unmarried man' for 'bachelor' in 'no bachelor is married'. All this discovery can do is to confirm empirically that there is indeed a relation of synonymy of some kind between the two expressions, but it cannot go beyond that and explain the presupposed notion of synonymy. It has no explanatory value. This notion of definition, conceived as correlating pre-existing relations of synonymy, is empirical and clearly can't explain analyticity — the latter requiring an elucidation that explains why this particular kind of truth is not based on empirical matters.[9]

Notice that in this account of the lexicographer's task, Quine attributes to him a notion of synonymy. If the lexicographer does not have a 'working' notion of synonymy, how can he confirm or discover that there is indeed a relation of synonymy between a pair of expressions? The interesting question here is to enquire into the nature of the lexicographer's notion of synonymy. If it turns out that the lexicographer has a clear and sharp notion of synonymy, it would mean that there is after all such a notion. The fact that this sharp notion of synonymy hasn't been given a satisfying theoretical characterisation would merely show that we haven't succeeded *yet* in finding one, not that the notion of synonymy — and meaning — is hopelessly unclear. Since Quine's argument in 'Two Dogmas' relies heavily on principled objections for the possibility of a satisfying clarification of the notion of synonymy, this point is of crucial importance.

In his discussion of the lexicographer's definitions, Quine himself hints at how

---

[9]In section 2, Quine goes on to examine variant kinds of definition, such as definitions used in formal work and definitions having an explanatory value. However, as Quine points out, they suffer from exactly the same problem: they turn out to rely on prior relations of synonymy. The only really transparent case of definitions creating synonymy relations are introductions of abbreviational notations, such as '$(a \neq b) =_{df} \sim(a = b)$'. But since this kind of definition has nothing to do with 'bachelor' and 'unmarried man', it can't be used to explain the analyticity of 'no bachelor is married' and the like.

this problem might be solved: "The notion of synonymy presupposed here has still to be clarified, presumably in terms relating to linguistic behavior" (1951: 24). This appeal to linguistic behaviour makes plausible the following interpretation of the lexicographer's notion of synonymy. In a given linguistic community, speakers often use terms interchangeably. For instance, it might be the case that when a speaker uses the word 'bachelor', he could as well have used instead the expression 'unmarried man'. This can be verified empirically, and we can picture the lexicographer's task as one of recording these correlations. But these observed correlations do not amount to an elucidation of the notion of sameness of meaning, since it is very unlikely that the empirical match between 'bachelor' and 'unmarried man' will be perfect. The working notion of synonymy used by the lexicographer is an approximate one, *i.e.*, a vague one. For him, two expressions are 'synonymous' if they are more or less used interchangeably by the speech community. His job is to report these vague correlations. Now if we take this notion of synonymy as the basis for an elucidation of analyticity, we should have to say that 'no bachelor is married' is analytic since 'bachelor' means more or less the same as 'unmarried man'. On this account, analytic truths depend both on meaning and empirical matter of facts, since the synonymy in play here is contingent.

Recently, Quine has re-addressed this issue and offered a slightly different perspective (1995a: 83). The lexicographer's task, he now tells us, is to specify the meaning of particular expressions. Since the lexicographer is after meanings, it seems that he must have both a notion of meaning, and a the notion of sameness of meaning — of synonymy. But Quine points out that the entries in a dictionary are there for explanatory purposes, exemplifying the use of the expression in typical circumstances. This is Quine's reason for claiming that the notion of meaning used by the lexicographer is not a theoretical one. The lexicographer's notion of meaning is to be understood in its vernacular sense. For Quine, there is no harm in talking of 'meaning' as long as

we do not engage in theorising about what meanings are: "I feel no reluctance toward refusing to admit meanings, for I do not thereby deny that words and statements are meaningful" (1948: 11). Quine is free to take the fact that sentences are meaningful as an irreducible one, or alternatively to seek an analysis of meaning in terms of behavioural criteria. The latter course is what Quine attempts in, for instance, chapter two of *Word and Object*. We will discuss these matters further in section 1.5.

1.2.2   Synonymy and necessity

In section 3, Quine returns to the problem of explaining synonymy and carefully examines a criterion designed for telling whether or not two expressions are synonymous. This attempt to capture the notion of synonymy is crucial. If indeed we can devise a clear and satisfying condition which will tell us, for any pair of words, whether they are, in the relevant sense needed here, synonymous or not, we will have solved the problem of explaining analyticity. The criterion under study here is the following. Two expressions are synonymous if and only if they are interchangeable *salva veritate*; that is to say, two expressions are synonymous if and only if putting one for the other in a true statement does not change the truth-value of that statement. It is hoped that this criterion gives us clear conditions of application for the notion of synonymy and therefore explains why statements like 'no bachelor is married' are analytic. Under this account, we could say that these statements are analytic since 'bachelor' and 'unmarried man' are everywhere interchangeable *salva veritate*, so that performing the substitution results in transforming 'no bachelor is married' into a logical truth.

As it will soon become clear, the force of the proposed criterion depends on the language in which it is applied. Quine starts by examining the criterion as it works in an intensional language and then studies its behaviour in an extensional language. For the sake of exposition, I'll reverse Quine's dialectic and study the criterion as it works in an extensional language first. For Quine, a language is extensional if "any

two predicates which agree extensionally (that is, are true of the same objects) are interchangeable *salva veritate*" (1951: 30). Now consider the obvious logical truth, couched extensionally:

(1) All and only bachelors are bachelors.

The needed interchangeability of 'bachelor' with 'unmarried man' is grounded on the following (extensional) statement:

(2) All and only bachelors are unmarried men.

What (2) says is that there is a *material* equivalence relation between the set of bachelors and the set of unmarried men: they consist of the same members. The statement (2) is our criterion for the synonymy of 'bachelor' and 'unmarried man'. It is a criterion for synonymy in the sense that if (2) is true, then 'bachelor' and 'unmarried man' are synonymous in the following sense: everything that is a bachelor is also an unmarried man. Assuming the truth of (2), we can produce from it the logical truth (1), therefore making a case that (2) is analytic since it is reducible to a logical truth by putting synonyms for synonyms. We can use the same routine to produce 'no unmarried man is married' from 'no bachelor is married' and (2), thereby showing, so the thought goes, that 'no bachelor is married' is analytic

The problem with this extensional approach can be illustrated by contrasting the example above with the following. Consider the logical truth:

(3) All and only creatures with a heart are creatures with a heart.

Consider also the following material equivalence, based on the model exemplified

earlier by (2):

(4) All and only creatures with a heart are creatures with kidneys.

On the extensional approach, (4) must be said to be an analytic statement since it is reducible to (3) by putting 'creatures with a heart' for 'creatures with kidneys' in (4). Since (3) is a logical truth, it follows that (4) ought to be analytic. It is clear, however, that the truth of (4) does not rest exclusively on the meaning of 'creatures with a heart' and 'creatures with kidneys'. It is an empirical (contingent) fact that these two expressions coincide in extension. Therefore, the truth of (4) does not rest exclusively on the meaning of its constituent expressions and fails to be analytic. In a nutshell, even if the expressions 'creatures with a heart' and 'creatures with kidneys' coincide in extension, they clearly don't have the same meaning,[10] showing that the extensional criterion for synonymy does not work. An extensional language simply does not have the means to devise a sufficiently fine-grained criterion for the purpose of explaining analyticity.

What is needed to explain analyticity is a criterion for what Quine calls *cognitive* synonymy, not mere extensional equivalence. Quine does not say much about the notion of cognitive synonymy, except that it is this kind of synonymy which is central for analyticity (1951: 28, 31). One way to interpret the notion of cognitive synonymy that would, I believe, be congenial to Quine is to stress that it is to be contrasted with synonymy understood as extensional equivalence. On this view, cognitive synonymy is simply non-extensional synonymy. If someone were to press the issue further and ask about the troublesome word 'cognitive', we could provide the following line of thought. Extensional equivalence is one kind of synonymy, but we have seen that it does nothing

---

[10]I assume, with Quine, that we are working with an intuitive notion of meaning and that the present inquiry is an attempt to make sense of that intuition.

to explain the notion of analyticity. This kind of synonymy is extensional in the sense that it is a function of how the world is: 'creatures with a heart' and 'creatures with kidneys' are extensionally synonymous due to the empirical fact that everything that is a creature with kidneys is also a creature with a heart. Concerning 'bachelor' and 'unmarried man', by contrast, we want to say that their synonymy goes beyond mere extensional equivalence: there is some yet unspecified connection between these two expressions, and a correct account of synonymy is after that unspecified connection.

A criterion for cognitive synonymy in the sense just explained can be devised by thinking about the difference between (2) and (4). Arguably, (2) is necessary whereas (4) is contingent. Is that the key for cognitive synonymy? That is to say, is *that* the unspecified connection? Extensional equivalence not being enough, we are led to the idea that these equivalencies must also be *necessary* to do the work and distinguish between the kind of equivalence needed for analyticity and mere extensional equivalence. On this account, to be cognitively synonymous — to have the same meaning — is to be necessarily interchangeable *salva veritate*. At first blush it seems to work. The statement:

(5) Necessarily, all and only bachelors are unmarried men

is true, whereas its counterpart 'necessarily, all and only creatures with a heart are creatures with kidneys' is not. If (5) is true, 'bachelor' and 'unmarried man' have the same meaning and it will form the basis to say that (2) is not merely true, but true by virtue of meaning — that is to say, that (2) is analytic. We appear to have explained why 'no bachelor is married' and the like are analytic: it is because their non logical expressions are cognitively synonymous, where the concept of 'cognitive synonymy' is to be explained in terms of statements involving the notion of necessity, such as (5). Our new criterion is fine-grained enough to discard unwanted exten-

sional equivalencies such as the one between 'creatures with a heart' and 'creatures with kidneys'.

But why is it that we are entitled to ascribe necessity to 'all and only bachelors are unmarried men'? The argument above presupposes that we are working within a non-extensional language. A non-extensional language is, for Quine, simply a language in which it is not the case that any two predicates which agree extensionally are interchangeable *salva veritate*. For instance, even if 'necessarily, all and only creatures with a heart are creatures with a heart' and 'all and only creatures with a heart are creatures with kidneys' are both true, the result of putting 'creatures with kidneys' for one of the occurrences of 'creatures with a heart' results in the falsity 'necessarily, all and only creatures with a heart are creatures with kidneys'. Here we have a clear-cut case of two predicates which agree extensionally yet are not interchangeable *salva veritate*. This is due to their occurrence within the scope of an intensional operator, in this case the operator of necessity. Such an intensional language, containing the concept of necessity, indeed affords the means to devise a suitable — that is to say, suitable for explaining analyticity — criterion for interchangeability *salva veritate*, as we saw above.

Let's grant that we all know that 'all and only bachelors are unmarried men' is true. Now what does it mean to say that it is not merely true, but necessarily so? Why does it have to be the case that all bachelors are unmarried men? The only available answer seems to be that all bachelors *must* be unmarried men because 'bachelor' *means* 'unmarried man'. If this is right, then the explanation is, again, circular. Using the notion of necessity was supposed to explain synonymy — which in turn was expected to explain meaning — but on close examination it turns out that synonymy and meaning were taken for granted all along.

It is important to point out that the notion of necessity in play here is applicable only to analytic statements, and that this is the reason why the approach fails in the

end. We have seen, in effect, that attaching the necessity operator to an extensional statement of equivalence results in a truth only when both sides of the equivalence have the same meaning, *i.e.*, when the statement of equivalence is not merely true but also analytic. This notion of necessity corresponds to Quine's first grade of modal involvement.[11] Statements, not propositions, are analytic, and necessity of the first grade consists of a semantical predicate attachable to names of statements. This is not a problem for Quine as long as we attach this semantical predicate to the first class of analytic truths, the logical truths. This class of truths is perfectly determined, so that there is no hesitation or vagueness as to the application of the necessity predicate over this class if we wish to treat logical truths as necessary truths.[12] But when it comes to ascribe necessity to analytic statements of the second group — those true by virtue of meaning — its conditions of applications are by no means clear because its applicability depends on relations of synonymy which are not sharply defined. Thus, says Quine, "the synonymy relation on which such cases [analytic truths of the second group] depend is supposedly a narrower relation than that of the mere coextensiveness of terms, *and it is not known to be amenable to any satisfactory analysis*" (1953a: 171, italics are mine).

In view of these considerations, it appears that we have two closely related arguments against the criterion of interchangeability *salva veritate*. The first is that we cannot explain analyticity with this criterion since the notion of sameness of meaning, and therefore of analyticity, is presupposed in the formulation of the criterion. This first argument is the one Quine is mainly concerned with in 'Two Dogmas', and is the one we have tried to make as clear as possible in the last few paragraphs.

The second argument is that the criterion does not work. We are, according to

---

[11]See Quine (1953c) for his view on the 'three grades' of modality.

[12]But such a notion of necessity is very 'meagre' for Quine. It does not even correspond to the notion of necessity as used in propositional modal logic (1953c: 171). This notion of necessity is merely to be equated with validity — that is, with one form of logical necessity as explained in the Introduction. We'll return to Quine's 'acceptable' notion of logical necessity in Chapter 4.

Quine, likely to hesitate over its application since it depends on a notion of synonymy which is not sharply defined. It is somehow assumed in 'Two dogmas' that synonymy, meaning and analyticity are not sharp notions. This is precisely why explaining analyticity in terms of synonymy fails: if it were possible to devise a sharp notion of synonymy, Quine's high standards would be satisfied, independently of the circularity in the explanation. In other words — to anticipate what follows — if it were possible to devise a satisfying identity principle for meanings, Quine's argument would collapse. The view that unsharp or somehow 'vague' notions are to be discarded or regimented can be seen as an ideal of logical positivism. This ideal, even if it is not defended and stated in 'Two dogmas', seems to play a crucial role in the dialectic of this paper, especially in the first three sections. On close examination, Quine is in effect arguing that the problem with the notion of analyticity is that we can't reduce it to or explain it in terms of other notions which are satisfyingly sharp. We will see in section 1.3, where we discuss criticisms addressing Quine's attack, how this second argument plays a crucial role.

1.2.3 Semantical rules

Section 4 of 'Two dogmas' examines the Carnapian proposal to use the notion of semantical rules as a key to explain analyticity. This part of Quine's paper is important since it has often been said that the problem with analyticity results from the vagueness inherent in natural languages. If we could make sense of analyticity in artificial languages, we would thereby show that the problems in explaining analyticity are due to the fact that we are dealing with unregimented, undisciplined natural languages. To make sense of a notion in the framework of an artificial language is, for Carnap and for Quine, to show its use where it counts, *i.e.*, in scientific discourse.

Semantical rules are part of artificial languages free of vagueness. The idea of such rules is to determine exactly (perhaps recursively) which statements are to be

analytic relative to a language L. These rules are semantical since they deal with *truths* of a particular sort, *i.e.* the analytic ones. Let's say we're working within a given artificial language L_0 — the details of that language are irrelevant. The role of the semantical rules is to specify exactly which statements of L_0 are to count as analytic. We thereby avoid the problem encountered more than once previously, namely the 'sharpness' problem: given these rules, we know exactly which statements are analytic and which ones are not, so that we never hesitate over the application of the term 'analytic' relative to that particular language. But from the point of view of explaining what is meant by the notion of analyticity, the appeal to semantical rules so conceived is misguided. The rules presumably have this form: 'a statement S is analytic for L_0 if and only if ___', where '___' specifies recursively a set of statements, which will consist of a subset of the true statements of L_0. But this only tells us which statements are to count as analytic in this language, they do not tell us *why* they are analytic. Indeed, if one were to ask why the statements specified by the rules are analytic, the sole response available (that is to say, the sole response not presupposing analyticity) would be that they are analytic because the semantical rules say they are. This obviously won't do: what we are after is an explanation, not just a list of analytic statements. It is the sole possible response in this context since any other explanation would have to involve the thought that the semantical rules determine the 'right' statements since they are so construed as to deliver all and only those statements which are true by virtue of meaning, thereby refocusing the problem on what is meant by 'true by virtue of meaning' and its cognates. Quine concludes: "Semantical rules determining the analytic statements of an artificial language are of interest only in so far as we understand the notion of analyticity; they are of no help in gaining that understanding" (1951: 36).

1.2.4 Summing up

We have seen that Quine has been refuting various attempts to explain the notion of analyticity. He used two main strategies to dismiss the explanations on offer. One is to show that the explanation does not clarify the notion, and the other is to show how the purported explanation depends on the notion of analyticity. In the remaining two sections, he addresses the second dogma and sketches his holist epistemology. These sections also contain a sceptical argument about analyticity and logical necessity, and we'll examine it in Chapters 2 and 3.

Now that we have rehearsed Quine's argument, let us ask: what is the relation between analyticity and necessity, and in which sense does the argument affect logical necessity? Quine is concerned with analyticity in Carnap's sense. For Carnap, analyticity coincides with logical truth, logical consequence, and validity. In other words, Carnap takes it that what is logically necessary in the sense explained in the Introduction will also be analytic. He is in fact, especially in his *Logical Syntax of Language*, trying to justify the concept of logical necessity through the notion of analyticity. A statement of the form $\Box(\Gamma \vdash_L \Theta)$ will be true, for Carnap, by virtue of the syntax — the rules — governing the expressions it contains. Such statements are *L-true* for Carnap, that is, analytic. Since Quine's attack is directed against Carnap's notion of analyticity, it follows that it is also directed to logical necessity. Quine's apparent conflation of the notions of analyticity and necessity is to be understood by relation to Carnap — he also conflates them, at least to a certain extent. The statements of logical necessity with which we are concerned here also happens to be analytic in Carnap's system, and Quine's main goal in 'Two Dogmas' is precisely to attack Carnap's notion of analyticity.

1.3 Grice & Strawson's defence

The first section was devoted to presenting Quine's argument against analyticity. We saw that his main criticism is that analyticity can't be satisfyingly explained. The problem is, roughly, that explaining analyticity by using synonymy or meaning fails since we don't have a clear (sharp, not vague) criterion of application for these notions. Grice and Strawson, in their joint paper 'In Defence of a Dogma', mount an elaborated and sophisticated defence against Quine's arguments in 'Two Dogmas'. We can discern two main strands in Grice & Strawson's section of the paper where they discuss Quine's argument as displayed in the first four sections of 'Two Dogmas'. The first line of defence concentrates on Quine's demands on explanation, that is to say, on Quine's idea of a satisfactory elucidation of a given notion. The second strand has to do with the implicit assumption made by Quine to the effect that the unintelligibility of a given notion follows from the fact that we haven't yet provided an adequate explanation for it. We will begin our investigation with the first counter-argument, *i.e.*, the claim by Grice & Strawson that Quine's demands on explanation are excessive.

Grice and Strawson contend that Quine raises the usual standards on explanation of concepts in 'Two Dogmas', and that this is why he condemns the notions of analyticity and meaning. If it could be shown that Quine is indeed raising the standard to an inappropriate level, it would follow that his attack on analyticity fails to be convincing. In this section, it will be argued that Quine's demands in 'Two Dogmas' are on the contrary perfectly reasonable ones. First, we will examine and criticise Grice and Strawson's interpretation of Quine's demands on explanation. Secondly, we will offer our own interpretation of what Quine's standards are. Finally, we will study the question whether these demands are reasonable or not and conclude that they are. Since they are reasonable and since Quine has shown that they are not satisfied, it follows that he is justified in maintaining that there is as yet no satisfying

explanation of analyticity — in a sense to be clarified. The further claim that the unintelligibility of analyticity follows from this lack of a satisfying explanation is of course another topic, taken up in the next section.

A good way to start the discussion is to quote the crucial passage in which they give their interpretation of what Quine's demands on explanations are:

> To make satisfactory sense (for Quine) of one of these expressions would seem to involve two things. (1) It would seem to involve providing an explanation which does not incorporate any expression belonging to the family-circle. (2) It would seem that the explanation provided must be of the same general character as those rejected explanations which do incorporate members of the family-circle (*i.e.*, it must specify some feature common and peculiar to all cases to which, for example, the word "analytic" is to be applied; it must have the same general form as an explanation beginning, "a statement is analytic if and only if... ").
> (1956: 147–48)

Let's begin by focusing on the first condition. According to Grice & Strawson, one of Quine's necessary conditions for explaining analyticity is that the explanation must not involve any of the terms belonging to the family-circle. Now what exactly is this circle? The first thing to notice is that Grice & Strawson seem to think (1956: 147) that the 'circle' in question is composed of *all* the candidate notions to explain analyticity examined by Quine in his paper. These notions, as we saw in the first section, are meaning, synonymy, interchangeability *salva veritate*, truth by definition, necessity, truth in all possible worlds, and semantical rules. Grice and Strawson's interpretation of Quine's view on this particular point presupposes that Quine's main contention is that analyticity is part of a circle consisting of these notions. However, I will offer two reasons against characterising Quine's main criticism in this way.

First, we saw in section 1.2 that the structure of Quine's argument is not, strictly speaking, a circle. For one thing, the attempts to explain analyticity by appealing to

semantical rules and possible worlds fail because the correct use of these supposedly explanatory notions relies on a prior understanding of analyticity. That is to say, in order to know that a sentence is true in all possible worlds, we must already know whether or not it is analytic. Concerning semantical rules à la Carnap, all they give us is a list of analytic statements; therefore, in order to have the 'right' rules, we must ensure that these rules will yield all and only analytic statements. But to do that, of course, we must already know what are the analytic statements. These explanations fail, not because the notions they use are part of a circle, but rather because the understanding of the notion we are trying to explain is presupposed. To make that clear, think about the following putative explanation. Let's say I want to explain to a child the concept of blue in the following way: 'something is blue if, and only if, when you see that thing, it is blue'. The criterion as such is not incorrect; it's just that it is of no use to understand the concept on the left-hand side — something more is needed, perhaps an ostensive definition in this particular case.

Explanations of this kind can be usefully compared to Dummett's 'pragmatic justifications'.[13] Such a justification, concerning logical laws in Dummett's context, proceeds by using the very logical law being justified. It certainly shows something about the law, and it can certainly be illuminating. But it cannot possibly be *persuasive* for someone who does not already consider as valid the logical law being justified. A persuasive justification of a logical law, it can plausibly be argued, must not employ the law being justified. Likewise, we could borrow Dummett's idea and distinguish between a pragmatic explanation and a persuasive explanation. The above explanation of 'blue' would be, on that account, a pragmatic explanation. Now the explanations in terms of semantical rules and possible worlds might be a bit more sophisticated, but they suffer from basically the same problem: in order to apply correctly the criterion, we must already know which statements are analytic. Such explanations are merely

---

[13]See Dummett (1991: 200–204).

pragmatic: they will only convince one who already has an idea of what analyticity is. It is clear that Quine's investigation is directed towards a persuasive explanation. He wants to show that one who is sceptical about the intelligibility of a notion cannot be convinced or persuaded that it is meaningful.

Strictly speaking, then, these notions (semantical rules and possible worlds) are not part of a 'circle' of explanation. Contrary to the other notions examined by Quine, the problem with them is not that they are unclear or lack sharpness, but rather that they *directly* presuppose the notion we are trying to explain. By 'directly', we mean that the intelligibility of the notion that is supposed to explain analyticity rests on the intelligibility of the notion of analyticity. It seems to be good reason to reject explanations of analyticity using this kind of strategy — indeed to reject any such putative explanation. In fact, Carnap himself never claimed that semantical rules could explain our concept of analyticity. They were rather meant to supply a precise characterisation, to 'precisify', as it were, our everyday concept of analyticity.

The second reason why it is not correct to picture Quine's main argument the way Grice & Strawson do is related to the circularity problem. If there is a circle in Quine's argument, it is composed of analyticity, meaning, synonymy, truth by definition, necessity, and back to analyticity again. We saw in the previous section that in order to explain analyticity, we might be tempted to use the notion of meaning. But this notion being unclear, we try explaining meaning by using synonymy. Synonymy not being sharp, we then try to explain it using truth by definition, but it turns out that definitions rely on a prior understanding of synonymy. Now if we try to explain synonymy by using necessity, we find that this latter notion presupposes analyticity. This looks more like an explanatory circle. The circle results from the fact that in the quest for an explanation of analyticity, we encounter unclear notions. This being so, we are forced to try to sharpen these notions by using other notions, but these latter notions turn out to be as unclear as the previous ones. We come full circle when we

attempt to elucidate the notion of synonymy by using necessity: Quine shows that in order to know whether a statement is necessary or not, we must know whether it is analytic or not.

In order to see that the circle is not, for Quine, the whole story, we can imagine what would have happened if the attempt to explain synonymy by using necessity was either not available or else resulted in a vague right-hand side. We would not have, strictly speaking, a circle, but rather a *chain* of unsatisfying explanations. Even if there would be no circle of explanation in this case, Quine would still claim that analyticity has not been satisfyingly explained. That is the important point: the explanations on the table are not clear. If on the other hand a case could be made that explaining analyticity in terms of, say, synonymy resulted in a clear and sharp elucidation, Quine would be forced to acknowledge the intelligibility of analyticity, circle or not. The circle results from one last attempt to explain synonymy (*i.e.*, the one using necessity), and it turns out that this last attempt presupposes directly an understanding of the notion of analyticity: the intelligibility of the notion of necessity (as construed by Quine) rests on the intelligibility of analyticity. The way I see it, circularity is a mere by-product of the whole attempt to elucidate analyticity.

The above considerations show why it is a mistake to interpret Quine as demanding, for a satisfactory explanation of analyticity, that it should not involve any notion which is part of the circle. If, for instance, we could provide an explanation of analyticity in terms of a sharp and clear notion of synonymy, Quine would have to be satisfied. It is true that if there were a satisfying explanation of analyticity by using synonymy, we would get out of the circle. But the primary goal is not to break out of the circle. It is, rather, to supply a clear characterisation. The first demand on explanation attributed to Quine by Grice & Strawson — namely the demand that analyticity should not be explained by using a member of the circle — is therefore misleading. It is misleading since it obscures what Quine really is after.

What about the second constraint on explanation attributed to Quine? Recall that it is the demand that the explanation should have the form 'a statement is analytic if and only if _____', where '_____' states sufficient and necessary conditions for a statement to be analytic. This way of interpreting Quine is faithful to the spirit of 'Two Dogmas', but with the proviso that '_____' *may* contain notions that are part of the explanatory circle, for the reasons explained above. This slightly modified requirement seems to be a *prima facie* reasonable demand on explanation. What is wrong, indeed, in requiring such explanations when one is doing philosophical analysis?

Let's take stock. Grice & Strawson contend that Quine's demands on explanation are unrealistic. They go as far as claiming that "it is perhaps dubious whether *any* such explanations can *ever* be given" (1956: 148). We have to keep in mind that given the way they represent Quine's exigencies, they might be justified in claiming that they are excessive. It is indeed difficult to imagine what an explanation of analyticity would be like if we could not use intensional notions such as meaning and synonymy. But Quine is not saying that there is a principled reason why we can't explain analyticity in terms of intensional notions: all he's asking, in the end, is that the notion(s) used be clear and sharp. If it turns out that there is a sharp notion of, say, synonymy, then Quine's demands would be satisfied. To repeat, it is crucial to keep in mind that Quine's main point in 'Two Dogmas' is *not* that analyticity is part of an explanatory circle, but rather that the explanations on offer do not *clarify* it. The circle results from the lack of a clear notion to explain analyticity — it is a by-product which does not play a crucial part in the argument. Since Grice & Strawson's interpretation of Quine's requirements is not fully faithful to what Quine actually maintains, they cannot, on the basis of that interpretation, accuse him of raising the standards.

But what, exactly, is Quine asking? We know he's after a clear and sharp expla-

nation, and now it is time to examine what Quine means by that. The most obvious and straightforward way to explain analyticity is by using the notion of meaning, since to be analytic is to be true by virtue of meaning. Remember that the friend of analyticity wants to say that 'all unmarried men are bachelors' owes its analyticity to the fact that 'unmarried man' and 'bachelor' have the same meaning. Now if we could tell, for any pair of expressions, whether they have the same meaning or not, then we would be equipped with a clear-cut criterion for sameness of meaning — think here of his famous slogan, 'no entity without identity'. Thus, we would also have a clear-cut criterion for analyticity, and Quine could not conclude that analyticity is unintelligible. This explains why the endeavour to elucidate analyticity in terms of meaning is crucial in 'Two Dogmas'. In effect, the attempt to explain meaning in terms of synonymy is also part of the project of explaining analyticity in terms of meaning, and so is the attempt to define synonymy by using truth by definition.

To sum up, Quine, working under the assumption (largely shared in the fifties) that meanings in the relevant sense are abstract entities, requires an appropriate principle of identity for meanings before accepting them into his ontology. This doesn't seem to be an extravagant demand at all. In the section following the next, I will develop in more details the role of the principle of identity in Quine's attack on analyticity. For the moment, let's turn to Grice & Strawson's second defence.

1.4 Does lack of satisfying explanation imply unintelligibility?

It may well be that analyticity is a concept poorly explained, that we don't have a satisfying account of it. The friend of analyticity can bite the bullet and claim that this state of affair has nothing to do with the intelligibility of the concept. The inference from the unavailability of a satisfying explanation for a given concept to its

unintelligibility is, she can argue, unjustified in the case of analyticity. In this section, we will examine carefully this line of thought, and see how a Quinean can reply to that objection.

Grice and Strawson, as well as Wright (1986), hold that there ought to be two steps in Quine's argument for Radical Irrealism about analyticity. The first step is to show that analyticity is inextricably entangled in an explanatory circle of concepts, or, as I prefer to interpret him, to show that analyticity is hopelessly unclear. The second step is that this unclarity, or circle, calls the meaningfulness of the notions involved into question. Quine persuasively argues for the first step, but virtually says nothing to buttress the second. This gap has been exploited by defenders of the analytic / synthetic distinction such as Grice & Strawson (1956) and Wright (1986). The substance of their counter-argument is nicely put in the following passage (Wright, 1986: 190)

> Quine, like Socrates, seems to have supposed that the absence of any clear, non-circular definition of a concept somehow calls its propriety into question. The proper response is that it does nothing of the sort, provided there is independent evidence that the concept is teachable and is generally well-understood. A sceptic about the intelligibility of a concept does not have to be answered by a rigourous explanation of it; it is enough to supply unmistakable evidence that the concept *is* well understood.

Wright is here explicitly blocking the inference from step one to step two: explanatory circle / unclarity do not entail the unintelligibility of the concepts involved. It is undeniable that Quine never discusses the implicit step uncovered by Wright. He is content to point out the lack of a satisfying explanation for analyticity, and leaves it at that. What we have to evaluate here is the force of Wright's consideration against a Quinean. To get started, let's have a look at Wright's criterion for deciding whether a concept is well understood, appearing immediately after the passage quoted above:

And what better evidence could there be than that there is manifest in the community at large a disposition towards non-collusive assent in the application of these concepts — crucially, in our application of them to previously unconsidered cases?

The key idea is that of non-collusive assent. An expression prompts non-collusive assent if a suitable community of speakers apply the expression in a more or less uniform manner without consulting each other as to the correctness or incorrectness of its applications. Wright's claim, then, is that the expression 'is analytic' generally prompts non-collusive assent in the philosophically informed community. That is, most of us ascribe analyticity to more or less the same cases, and that very fact suffices to show, according to Wright, that the concept of analyticity is well understood. Therefore, the argument goes, Quine is wrong to say that the concept of analyticity is unintelligible.

Let us grant, if only for the sake of the argument, that analyticity *does* prompt non-collusive assent.[14] This line of thought presents itself as a real threat to the Quinean for the following reason. The criterion of non-collusive assent is congenial to Quine, since it is empirically observable whether an expression satisfies it or not. He can therefore not afford to dismiss the criterion on the basis that it is not sufficiently clear, ill defined, vague, etc. Quine can only agree that — if indeed it is the case — 'is analytic' prompts non-collusive assent. At first blush, then, if Wright is correct to hold that non-collusive assent implies intelligibility, Quine would be refuted.

One possible reply available to the Quinean is to resist the thesis according to which non-collusive assent implies intelligibility. One of Quine's criteria for intelligibility is, we have seen, clarity in the explanation of the concept. Wright dismisses this

---

[14]It's not clear that it does. Any logical law, the paradigm case of analytic statements, can be disputed. That is, for any logical law, there is at the very least a community of speakers who would not apply the expression 'is analytic' to it. Hume's Principle is another disputed case, and so is 'everything green is extended' (this last is taken from Quine (1951: 32)).

criterion and suggests another one, non-collusive assent. Quine could thus challenge Wright's criterion and put forward arguments to the effect that his criterion is better than Wright's, or that, simply, Wright's criterion is not acceptable for such and such reasons. This strategy is, I take it, very risky. It would mean that Quine has to show that despite the fact that speakers generally agree on the application of a given expression, it remains possible that these speakers are just talking nonsense every time they do. This is very counter-intuitive — the arguments to motivate this view would have to be correspondingly complicated and contentious.

I have no intention, however, to defend that view. On the contrary, I want to suggest that a Quinean can: a) accept Wright's thesis about the link between non-collusive assent and intelligibility and b) accept that the expression 'is analytic' prompts non-collusive assent, but c) nevertheless stick to her Radical Irrealism about analyticity. The key is to reflect on what exactly satisfaction of Wright's criterion for intelligibility establishes for a given expression. The most such a criterion can establish for analyticity is that there is, corresponding to the expression, *a* concept of analyticity used by the speakers of a suitable community.[15] In other words, it establishes that speakers are not uttering sheer nonsense when applying the expression 'is analytic' to statements.

Does that signal defeat for Radical irrealism? I wish to offer a reason why I think it does not. Competent speakers mostly agree on the application of the concept of analyticity because they mostly agree on their use of the concept of synonymy. Speakers uniformly say that 'all bachelors are married men' is analytic in view of the fact that they agree that 'bachelor' and 'unmarried man' are synonymous expressions. They're going to say, if pressed, that the reason why they ascribe analyticity to this statement is because the two expressions involved mean the same thing, and that this

---

[15]Grice and Strawson claim that non-collusive assent does even more than that. According to them, non-collusive assent, in the case of analyticity, means that there is, in a realist sense, an analytic / synthetic distinction.

relation of synonymy is why, ultimately, the sentence is true by virtue of meaning. So it is beyond doubt that competent speakers have a concept of analyticity. Their concept is the expected one: for them, an analytic statement is one true by virtue of meaning alone. Underlying this uniform use of the concept of analyticity is an agreed understanding of synonymy and of which expressions are synonymous with each other.

However, this does not show that competent speakers are endowed with the concept of analyticity that Quine has in mind. Indeed, as we will see in the next section, Quine's target is not the intelligibility of analyticity qua notion, but rather its Platonistic construal. What Wright's argument shows is that competent speakers have *a* concept of analyticity — nothing about the nature of the concept is shown. Quine would be happy to agree with Wright and he would certainly suggest that analyticity has to be understood in terms of assent and dissent, which is just what he does in section 14 of *Word and Object*. But such a construal of analyticity need not commit one to the existence of a realm of a priori knowledge. Since the a priori is, as I will argue in the next section, Quine's main target in 'Two Dogmas' and elsewhere, Wright's argument perhaps fails to address the heart of the controversy.

1.5 Boghossian's *reductio*

After the publication of Quine's 'Two Dogmas of Empiricism', the notion of analyticity fell into disrepute for several decades. It is only recently, thanks to the work of Boghossian and others, that analyticity started to regain a respectable status in the philosophical world. In his paper 'Analyticity',[16] Boghossian defends the notion against Quine's forceful and influential attack. One contention of 'Two Dogmas' is that there is no fact of the matter as to whether an expression is analytic or not. In

---

[16] All page references in this section are to this paper, unless specified otherwise.

other words, Quine maintains a non-factualist thesis about analyticity. Boghossian, in the aforementioned paper, puts forward an argument to show that *non-factualism* about analyticity entails scepticism about meaning — and this very fact supposedly shows that Quine's argument can't be right. In sum, Boghossian's argument is intended as a *reductio ad absurdum* of non-factualism about analyticity. In this section, I present an argument to the effect that Boghossian's attempted *reductio* fails against the Quinean. In the course of making this case, I will suggest that Quine's argument, properly interpreted, is successful as it stands but has a more limited scope than is generally believed.

1.5.1 Quine's argument

Quine's charge against analyticity in 'Two Dogmas' mainly proceeds through a criticism of the notion of meaning. The reason for this is straightforward. Quine is interested in the notion of analyticity in its contemporary sense: to be analytic, in this version, is to be true by virtue of meaning. In particular, he wants to attack the notion of analyticity as it appears in the writings of Carnap, *e.g.*, in *The Logical Syntax of Language* and *Meaning and Necessity*. Analyticity plays a pivotal role in these two books, so much so that a definitive rebuke of the notion would mean that the foundation of Carnap's methodological framework is flawed at its very core. Carnap's notion of analyticity, of course, is inextricably tied to meaning. Thus Quine, by mounting his criticism of analyticity on the back of meaning, threatens to demolish the underpinnings of Carnap's system.

Let's have a closer look at the structure of Quine's argument. A sentence such as 'all bachelors are unmarried men' is analytic, according to the proponents of the notion, due to the fact that 'bachelor' and 'unmarried man' mean the same thing. This is precisely what Quine wishes to undermine by arguing that we don't have an adequate explanation of what it is for two expressions to mean the same thing in the

relevant sense. Therefore, the argument goes, we don't have an adequate explanation of analyticity. Since such an explanation is neither available nor forthcoming, it follows that analyticity is not a factual notion. We can reconstruct the argument as follows:

1. the notion of analyticity is best cashed out in terms of the notion of 'meaning the same as'

2. there is no satisfying explanation of the notion of 'meaning the same as'

3. if there is no satisfying explanation of a particular notion, then there is no fact of the matter about its application

4. therefore, we should be non-factualists about the notion of 'meaning the same as'

5. since (1), non-factualism about analyticity follows

As seen in the previous two sections, there is more than one way to dispute the correctness of this argument. Grice & Strawson's (1956) well-known criticism of Quine's argument focuses on (2) and (3). Wright (1986) adopts a similar strategy. According to them, it is not clear that there is no satisfying explanation of analyticity (or of 'meaning the same as'). Furthermore, they argue, even if Quine is right in this respect, it does not follow that the notion is not factual. That is, they dispute both (2) and (3).

However, I wish to examine another counter-argument, that put forward by Boghossian. His strategy is to show that (4) leads to an absurdity. That is, Boghossian does not try to convince us that there is a satisfying explanation of 'meaning the same as' by supplying one: he claims rather that to hold that there is no such explanation is self-refuting. Grice & Strawson anticipated an argument of this kind in their 'In Defence of a Dogma':

We want only to point out that if we are to give up the notion of sentence-synonymy as senseless, we must give up the notion of sentence-significance (of a sentence having meaning) as senseless too. But then perhaps we might as well give up the notion of sense. — It seems clear that we have here a typical example of a philosopher's paradox. (Grice & Strawson, 1956: 146-147).

If Grice & Strawson's line of thought is persuasively exploited by Boghossian — if, that is, the crucial premise (4) is false — it follows that Quine's argument against analyticity does not go through. The advantage of such an approach is that it dispenses with the need to invoke the battery of complicated arguments in order to dispute (2) and (3). Unfortunately, as we will see, the *reductio* is fallacious.

1.5.2   Boghossian's anti-Quine argument

Before we examine Boghossian's argument, we need to put it into context. In particular, we have to distinguish between three notions of analyticity. In his 'Analyticity', Boghossian puts forward a notion of analyticity — which he calls 'epistemological analyticity' — in the hope that it will provide, in his words, an 'explanation of the a priori' (p. 332). Epistemological analyticity is defined thus (where 'S' refers to an arbitrary sentence):

Epistemological analyticity: S is analytic if grasp of the meaning of S suffices for holding S true. (See p. 334).

This notion is to be contrasted with metaphysical analyticity:

Metaphysical analyticity: S is analytic if it owes its truth-value completely to its meaning, and not at all to the 'facts'. (*Ibid.*)

For a set of reasons that need not be discussed here, Boghossian is not interested in the metaphysical notion and indeed praises Quine for making us realise its inappro-

priateness (p. 335). In the section discussed here, Boghossian is concerned to defend a type of epistemological analyticity, namely what he terms 'Frege-analyticity':

Frege-analyticity: S is Frege-analytic if it is transformable into a logical truth by substituting synonyms for synonyms. (p. 337).

This notion of analyticity, as is well known, is the one attacked by Quine in 'Two Dogmas'. Indeed, in that paper, Quine discusses neither the analyticity nor the putative a priority of the logical truths. But a considerable set of analytic statements fall into the category of Frege-analyticity; so if sense can be made of their analyticity, then the problem of the a priori will be reduced to that of making sense of the analyticity of logic (as Boghossian himself points out (p. 339)). Boghossian's hope is that a vindication of Frege-analyticity — a vindication, namely, that sentences such as 'All bachelors are unmarried men' are analytic — will provide the means to defend the claim that such statements are known a priori. It is important to keep in mind that Boghossian's ultimate aim is to defend a priori knowledge and that Quine's aim, as we will discuss below, is to undermine this kind of knowledge.

Let us scrutinise in more details the attempted *reductio*. The argument levelled against Quine is that to claim, as he does, that there is no satisfying explanation of the expression 'meaning the same as' leads to an outright rejection of the notion of meaning itself. And this, Quine's opponents argue, is totally unacceptable. According to Strawson & Grice, as we have seen, it shows that Quine's line of thought leads to a 'paradox', whilst for Boghossian (p. 342) it amounts to a *reductio* of his rejection of analyticity.

Boghossian (pp. 342–43) gives a compelling argument to support his claim. How is it possible, he maintains, that there be a fact of the matter as to what each expression means without there being a fact of the matter as to whether two expressions mean the same? That is, given that some expression E determinately means M, how could it be indeterminate whether another expression F, which also possesses a determinate

meaning, does or does not mean the same as E? Given the two meanings, it should be a determinate matter whether they are the same or not. And since Quine argues that there are no facts of the matter concerning the synonymy of two expressions, he must, logically, endorse the thesis according to which there are no facts of the matter as to what each expression means. On examination, it appears that Boghossian is here using a contraposing argument. Take 'FM1' to mean 'there are facts of the matter as to what each expression means' and 'FM2' to mean 'there are facts of the matter as to whether two expressions mean the same or not'. Boghossian puts forward an argument to show that FM1 → FM2. Given the truth of FM1 → FM2, we just have to contrapose to get ~FM2 → ~FM1. This last conditional is what leads to the *reductio*: it shows that Quine's criticism of synonymy — FM2 — leads inevitably (logically) to a rejection of meaning — FM1. The argument can be formalised thus:

1. FM1 → FM2   Premise

2. ~FM2 → ~FM1   Contraposition on (1)

3. ~FM2   Quine's thesis ('4' in the argument above)

4. Therefore, ~FM1   Modus ponens on (2) and (3)

There is thus no doubt that Boghossian's reasoning is valid. Prima facie, it also seems that Quine should accept not only the validity of the argument, but also its soundness. Indeed, Quine repeatedly mentions that he dislikes the notion of meaning. It is explicit, in 'Two Dogmas', that he intends to show that the notion of meaning stands or falls with the notion of analyticity. In the attempt to explain analyticity, we have, he maintains, to use the notion of meaning. But the appeal to the notion of meaning does not help, since that notion is as unclear and hopeless as the notion of analyticity itself. Therefore, since he rejects analyticity on the ground that it does not satisfy his standards of explanation and since meaning likewise does not satisfy

these standards, he would be inconsistent if he tried to discard analyticity without discarding, at least in some sense, meaning as well. In fact, this is a consequence of the 'domino effect' inherent in the structure of Quine's argument. His criticism of analyticity has consequences, if he is right, for all the notions in the explanatory circle, which consists, principally, of the notions of analyticity, meaning, necessity, and synonymy. In this sense, Boghossian's argument shows something that Quine is perfectly aware of and is glad to accept. If he is, how come he would not (presumably) consider Boghossian's argument as being a *reductio*? That is, why is the rejection of meaning seen as a destructive consequence for Quine's views according to Boghossian, but not for Quine himself? It is true indeed that the consequences of a complete rejection of meaning are difficult to imagine. If Quine really is jettisoning the notion of meaning — *any* notion of meaning — he owes us an explanation of how he intends to live without meaning. So we are, it seems, facing a dilemma: either Quine holds an incredible thesis about meaning, or there is something wrong with Boghossian's argument — despite all appearances to the contrary.

1.5.3   Exegetical study: Part I

Quine's ideas about meaning have often been taken to suggest that he is trying to get rid of the notion altogether.[17] And his writings strongly suggest that interpretation. But that reading is, I think, not wholly faithful to what Quine is after in his criticism of meaning. I want to suggest that Quine does not maintain that he can actually do without the notion of meaning as such, and moreover that there need be no inconsistency on his part. That is, there is conceptual space for both a criticism of the notion of meaning such that it fulfils the role Carnap intends it to have *and* the acceptance of meaning as a legitimate notion. I wish to argue that Quine occupies that conceptual space.

---

[17]See for instance Strawson (1985: 69-95). For a discussion of this topic, see Quinton (1990).

In order to drive that point home, we have to reflect upon Quine's arguments in support of the premise (4) in his argument for the non-factuality of analyticity. The first thing to bear in mind is that Quine assumes (for reasons to be explained shortly) that meanings are (abstract) objects. The relation of 'meaning the same as' is therefore a relation that holds or fails to hold between objects of a particular kind. The second thing to remember is how these considerations relate to Quine's dictum 'no entity without identity'. Since meanings are objects, they have to be provided with suitable criteria of identity if they are to be accepted in his ontology. Also, since 'meaning the same as' is a relation of identity, the attempt to explain satisfactorily this relation is crucial in two respects. Firstly, it is needed to make sense of the analyticity of statements such as 'all bachelors are unmarried men' in view of the fact that the putative analyticity of this kind of statements depends on the synonymy of 'unmarried man' and 'bachelor'. Secondly, it is needed to render intelligible the very notion of meaning. Quine is therefore acutely aware that (2) amounts to a rejection of meaning, given the fact that it follows according to standards of his own.

In the light of these preliminary considerations, what would be a satisfying elucidation of the notion of meaning for Quine? It must be recalled that Quine attacks meaning as it is needed for an appropriate notion of analyticity. In addition, we must bear in mind that such a notion derives its utility in its capacity to explain a priori knowledge. Indeed, this is recognised by Boghossian himself: ' "All bachelors are male" does seem to be (Frege-Analytic), and that fact does seem to have something important to do with that statement's apriority' (p. 339). To be sure, why else would we need a notion of analyticity if not to explain a priori knowledge? Arguably, this is the reason for its introduction in the history of philosophy. After it was recognised that explaining a priori knowledge via intuition is a project doomed to failure, and after the linguistic turn made possible the introduction of a notion of analyticity understood in terms of meaning, the project to explain the a priori was revived

by, amongst others, the logical positivists.[18] Another thing to bear in mind is that a suitable conception of meaning — one adequate for explaining the a priori — is the abstract, Platonist conception of meaning. We can see here two main strands in Quine's thought working hand in hand, namely his anti-Platonism concerning meaning and his anti-a priori outlook. The idea, roughly, is that if we somehow have access to an abstract realm of objects, the connection between these objects will be known a priori. Knowledge of meanings to access the realm of the a priori replaces intuition. Again, this is acknowledged by Boghossian: "Indeed, it seems to me that epistemic analyticity can be defended quite vigorously, especially against the background of a realism about meaning" (p. 358). Accordingly, Quine's target is meaning understood in that sense. Let us recapitulate the exegetical points just made:

- Quine attacks meaning as it is needed for an appropriate notion of analyticity.

- Such a notion derives its utility to explain a priori knowledge.

- An abstract, Platonistic notion of meaning does the job.

- Accordingly, Quine's criticism targets such a notion of meaning.

It is clear that if this is correct, Quine attacks only a particular conception of meaning, not the notion as such. This subsection and the next are devoted to defending this claim.

The key to understand Quine's target in 'Two Dogmas' lies in his well-known dictum 'no entity without identity'. Indeed, a satisfying elucidation of the Platonist notion of meaning must go through a specification of the identity condition for such objects. For Quine, not only is this a necessary condition for the acceptability of meanings; it is also sufficient (Quine, 1995a: 75–76). A careful reading of the crucial section 3 of 'Two Dogmas' reveals that this is exactly what Quine is after: a criterion

---

[18]See Coffa (1991) for a fascinating historical account of these issues.

of identity for meanings in terms of synonymy. A criterion of identity for meanings would have the form:

(**CIM**) $\forall\alpha\forall\beta(($the meaning of $\alpha =$ the meaning of $\beta) \leftrightarrow (\alpha \approx \beta)),$

where the Greek letters stand for names of expressions — say 'α' for ' "bachelor" ' and 'β' for ' "unmarried man" '— and '$\approx$' represents some equivalence relation. The essential point is that we can illuminate Quine's remonstrance about meanings by reading him as saying that there is no satisfying equivalence relation on the right-hand side of (CIM) to derive from it a criterion of identity for meanings. These considerations are directly relevant to the second premise in Quine's argument as explained above since (CIM) is in effect an attempt to explain what it is for two expressions to 'mean the same'. The most promising instance of (CIM) for meanings is the version in which meaning is defined with the help of synonymy:

(**CIM***) $\forall\alpha\forall\beta(($the meaning of $\alpha =$ the meaning of $\beta) \leftrightarrow (\alpha$ is synon. with $\beta))$

It is plausible, and fruitful, to read section 3 of 'Two Dogmas' as criticising the viability of (CIM*) on the ground that the relation of synonymy on the right-hand side can't do the job, *i.e.*, individuate meanings in the appropriate manner.

1.5.4 Exegetical study: Part II

We saw that Quine's views on meaning are best understood in connection with the identity principle. This principle affords a method for individuating objects (abstract or concrete), thereby making such objects 'acceptable' in one's ontology. Quine holds, under my interpretation, that there is no suitable identity principle for meaning. Therefore, the view goes, meanings can't be accepted in a properly regimented ontology. I take it that this is the crux of Quine's remonstrance about meaning: we can't make sense of meanings as (abstract) objects. This is what he establishes in 'Two Dogmas' and elsewhere. It is a compelling argument, as long as we don't

'inflate' it and read too much out of it. The argument establishes (convincingly, in my opinion) that Platonism about meaning is false. Platonism about meaning is the view, very roughly, that meanings are somehow 'out there' in the abstract world and that our words and sentences attach mysteriously to these abstract objects. What Quine shows is that meaning so construed is a myth: there is no such thing.[19] In this sense, Quine is attacking a particular conception of meaning, not the notion as such. The only way we can take him to criticise the very concept of meaning is to say that for him, the notion of meaning is irreducibly tied to Platonism. A criticism of Platonism would therefore amount, implausibly, to a rejection of meaning. A parallel with the philosophy of mathematics might help to put these considerations in perspective. Those who argue against Platonism in the philosophy of mathematics do not necessarily wish to deny that the concept of number is a legitimate notion — in fact, I can't think of one single philosopher who holds that view. What they reject is a particular conception of numbers, namely the view according to which numbers are abstract objects and that numerals refer to these objects.[20]

It is true that Quine sometimes gives the impression that he wants to reject the very notion of meaning. His paper 'The Problem of Meaning in Linguistics', written a few months after 'Two Dogmas', is further evidence for that interpretation (1953b: 47-49) but it also contains what purports to be accounts of synonymy and of significance. It is plausible to hold that in that paper, Quine is trying to reduce the concept of meaning to an acceptable notion of synonymy and also to the less problematic linguistic notion of significance. Talk of 'significance' instead of 'meaning' is recommended by Quine (1948: 11; 1953b: 48-49) on the ground that the notion of significance does not have the ontological connotations associated with the notion of

---

[19]Wittgenstein attacks the same conception in his *Philosophical Investigations*.

[20]It should be clear that Quine has nothing against abstract objects *per se*: it is not because meanings as objects are abstract that he does not accept them in his ontology, but rather because there is no criterion of identity for them. He is happy to accept sets and other well-defined abstract objects.

meaning.

What is perhaps misleading regarding Quine's attitude towards meaning is his insistence that we have to 'get rid' of intensional notions in general, including meaning. For Quine, it seems that to criticise the view according to which meanings are abstract objects is to criticise the very notion of meaning. This way of seeing the matter can be justified through the thesis of the priority of syntax,[21] since if talk of meaning does commit one to an ontology of meanings as abstract objects, and if, in addition, it can be conclusively shown — as Quine tries to do — that we can't make sense of such objects, then there is every reason to jettison the notion in question. The problem here is, of course, that the assumptions under which the argument depends are controversial, especially the one having to do with the priority of syntax. There seems to be every reason to think that the notion of meaning can be interpreted in a non-Platonistic way.

1.5.5   Boghossian's argument: The revenge

Equipped with this interpretation of Quine's views, it appears at first blush that we have a convincing case against Boghossian's attempted *reductio*. By restricting the scope of Quine's argument, it is tempting to conclude that it leaves open the possibility of a non-Platonistic construal of meaning, and as such the Quinean argument displayed in (1) – (5) would not be committed to meaning-scepticism. Unfortunately, things are not so simple since Boghossian draws a distinction between *monadic* and *relational* construals of meaning. A relational construal of meaning is one according to which any significant expression E bears a relation R to its meaning M. It is plain

---

[21]The thesis of the priority of syntax is the thesis according to which the appropriate syntactic category of expressions somehow determines their ontological import. Take for instance singular terms: if it can be shown that an expression belongs (irreducibly) to the syntactical category of singular terms, this is taken as strong evidence that the singular term in question *refers to an object*, given that the term appears in true sentences. Talk of *the* meaning of such and such expressions and of *having* the same meaning as such and such expressions, and so forth, commits us to the existence of meanings according to this thesis.

that such accounts of meaning would be unacceptable for a Quinean, since it involves talk of *the* meaning of given expressions, and such talk commits one to the ontology Quine is criticising. Any relational construal of meaning appeals to meanings as objects, since it depends on irreducible meaning-facts. Quine certainly does hold that we must avoid the notion of meaning so conceived, and in this sense he would regard Boghossian's argument not as a *reductio*, but as a valid and true argument. That is to say, Quine would be happy with Boghossian's conclusion, if all it involves is merely a rejection of meaning in this sense.

But Boghossian goes further and argues that his argument is stronger than that. According to Boghossian, the argument he offers against non-factualism about analyticity will go through for any 'monadic', as well as relational, conception of meaning.[22] This kind of construal of meaning has it that any meaningful expression E has a monadic property P which is central to the conception of meaning in question. For instance, in a behaviouristic conception, an expression E provokes the disposition D; the meaning of E is to be explained on the basis of that disposition. In a version of use-theoretic accounts of meaning, any meaningful expression E has a history of use H; this property H is used to explain the meaning of E. Other monadic conceptions follow the same pattern. A monadic conception of meaning is subjected to the *reductio* since any such account requires appeal to facts of the matter, *i.e.*, in order to explain the meaning of E we have to look at, for instance, the disposition assigned to it, which is, in principle, something determinate. It is because it is determinate — because there is a fact of the matter as to what an expression E means — that it is subjected to the argument.

In sum, to resume the thread of the central argument, it would appear that Boghossian holds the view according to which an attack on the notion of analyt-

---

[22]'The question naturally arises whether this result is forthcoming *only* against the background of a relational construal of meaning. I think it's quite clear that the answer is "no"' (p. 343).

icity through a criticism of the notion of meaning, either on a relational or monadic construal, is self-refuting. This argument of Boghossian is designed to show why Quine's attack on analyticity leads inevitably to a total rejection of meaning. So it would seem to mean that it doesn't matter if — as I have done a few paragraphs back — we interpret Quine as merely jettisoning the Platonistic conception of meaning. If Boghossian is right, Quine is in effect jettisoning any conception of meaning without realising it.

1.5.6 Quinean defence completed

If this were the case, it would truly be disastrous for Quine. How can a philosopher reject any explanation of meaning? It is clear that some expressions have sense, mean something, are significant, etc. An argument to the effect that any explanation of that phenomenon is impossible requires much more than what we can extract from 'Two Dogmas'. The truth is that Quine would not object, on principled grounds, to a monadic construal of meaning. He would certainly be happy, for instance, with a use-theoretic account (Quine, 1981). To see that, let's reflect on an important feature of monadic construals. They are, in effect, reductive accounts of meaning: they reduce meaning to something else, something that can, in the best cases, be readily and determinately identifiable. It is certainly in accordance with a Quinean spirit that reductive accounts are put forward. At the very least, such accounts are preferable to relational ones since they do not depend on a Platonistic conception of meaning.

In fact, Quine attempts such a reductive, monadic construal of meaning in Chapter two of *Word and Object*. The notion of 'meaning the same as' is to be reduced to that of 'stimulus-synonymy' and, in general, the meaning of an expression is to be understood in terms of the concept 'stimulus meaning', which is cashed out in terms of assent and dissent (1960: Chap. 2). This is, roughly speaking, an account of meaning

in terms of disposition: an expression's meaning is reduced to the set of dispositions it is likely to provoke in any given speaker. It is in this sense that such an account is monadic, and not relational. If Boghossian is right, Quine is contradicting himself when putting forward this theory of stimulus meaning, since according to Boghossian, Quine is committed — on the basis of his criticism of analyticity — to reject any account of meaning.

Putting aside this alleged contradiction for a moment, let's pause to reflect on the status of Quine's theory of 'meaning'. This theory fits nicely in Quine's general ideas about the nature of philosophy and about ontology. Concerning the first, recall that for Quine, philosophy is a branch of empirical science.[23] The notion of stimulus meaning is clearly one that is empirical: there are facts of the matter (as Boghossian would put it) as to whether an expression provokes this or that stimulus. Concerning Quine's ontological preferences, it almost goes without saying that stimulus meaning is a notion that accords particularly well with his strong taste for a 'purified' ontology, in which preference is to be given to objects somehow 'sanctioned' by empirical observations. In short, there is nothing distinctively troublesome in an ontology of stimuli since it is basically an observational notion: one can observe a stimulus.

Now, is Quine contradicting himself? Does his argument against analyticity as displayed in the first four sections of 'Two Dogmas' really commit him to reject any account of meaning, including the dispositional one outlined in *Word and Object*? Boghossian's argument seems, again, to trap Quine into a very uncomfortable dilemma: to avoid contradiction, it seems that he must either withdraw his criticism of analyticity or reject any account of meaning, be it relational or monadic. Both horns of the dilemma are non-options for Quine. As to the first, it is clear that his rejection of analyticity plays a pivotal role in his whole system, especially concerning epistemology and ontology, two central areas of his thought. The second is plainly

---

[23]His paper 'Epistemology Naturalized' is the classic exposition of this thesis (1969a)

absurd, as Boghossian points out (p. 342). To deny the possibility of any account of meaning seems analogous to denying any account of colours: it is undeniable that we see colours, and thus we are justified in asking for an explanation of that phenomenon. By the same token, since it is equally undeniable that expressions are meaningful, we are justified in asking for an explanation of that evident fact. It does not matter if, in the end, it turns out that there are no colours 'out there', or that there are, in the final analysis, no meanings in the world. What matters is an explanation of the phenomenon independently of the required ontology to explain it.

The dilemma, then, is fatal. But the truth is that Quine does not face it. His criticism of analyticity, as it stands and as I have interpreted it, is perfectly consistent with a monadic construal of meaning. It is true, however, that it is not consistent with a relational account of meaning, since such a construal requires an ontology of meanings. But Quine will be happy to reject relational accounts. If Quine is, as I argue, consistent, where does Boghossian's argument go wrong? Or, to put the question differently, in which sense does Quine's account of meaning not commit him to accepting the notion of analyticity? In order to get a grip on that crucial issue, it is necessary to step back a bit and reflect on the nature and target of 'Two Dogmas'. Quine's main target was the putative division between statements that are true by virtue of meaning alone and those that owe their truth, in addition to what they mean, to 'facts' (1951: 20). If there is such a division, it follows that there are truths that are somehow knowable a priori. These truths have a special status: they are not empirical. This supposedly non-empirical status is what Quine is determined, in the end, to vilify. His attack on analyticity is merely instrumental to achieving that goal. Logical positivists, most notably Carnap, 'cashed out' this a priori / a posteriori distinction in terms of the analytic / synthetic distinction (see Coffa 1991: 272-326). Therefore, in order to criticise the cleavage as understood by the logical positivists, Quine proceeds through a criticism of analyticity.

The key point is that Quine attacks the notion of analyticity only insofar as it is characterised in such a way as to be able to make sense of a priori knowledge. It is difficult to find an explicit admission of this in his own writings, but it locates the argument against analyticity in 'Two Dogmas' in a compelling perspective. Indeed, why would Quine be so keen to criticise the analytic / synthetic distinction if not to show the illusion of a priori knowledge? What is the value of analyticity *in itself*, without regards to its explanatory power in relation to a priori knowledge? Analyticity for analyticity's sake is useless. It derives its utility when one is attempting to explain a priori knowledge — Boghossian himself admits that, as we have seen above. However, I grant that this interpretational standpoint has to be taken as a working assumption. On the plus side, this — plausible — assumption affords the interpreter of Quine's philosophy with an enlightening viewpoint. In any case, even if it is hard to find an explicit quotation to the effect that Quine's attack on analyticity is a means to destroy the foundations of a priori knowledge, this interpretation is consistent with his generally anti-a priori outlook.

This is crucial since even Quine is ready to acknowledge a notion of analyticity built from stimulus meaning. That notion of analyticity might be intelligible, but it won't be able to provide a foundation to a priori knowledge. What such an account of analyticity does is merely to specify which sentences a speaker, or at best a set of speakers, assent to 'come what may' (to be understood as, perhaps clumsily, 'come what stimulation may' (Quine, 1960: 66)). The notion of stimulus analyticity might be intelligible, but it is not discriminating enough: indeed, on this account of analyticity, the sentence 'there have been black dogs' will come out analytic for most speakers. But this Quinean construction of analyticity, admittedly, relies on a conflation of analyticity with necessity.

At this point one could raise the following objection:

Quine shoots himself in the foot when he puts forward his dispositional account

of meaning. Indeed, the notion of stimulus-synonymy does provide us with a suitable right-hand side for (CIM). On this view, two expressions have the same meaning if and only if they produce the same pattern of assent and dissent. As such, 'bachelor' and 'unmarried man' do have the same dispositional pattern and so they have the same stimulus-meaning. Therefore, the Quinean model provides us with a model for the analyticity of sentences such as 'all bachelors are unmarried men'.

However, this objection betrays a misunderstanding of the scope of the notions of stimulus meaning and stimulus-synonymy. Stimulus meaning has to be relativised, at best, to a speaker and to a slice of time (the so-called 'modulo'). It is hard to see how we can get a notion of meaning that is appropriate to explain a priori knowledge from such a limited notion. According to Quine, the closest we can get to a notion of analyticity (one not conflated with necessity) with stimulus meaning is an externalised notion:

> In *Roots of Reference* (pp. 70-80) I suggested externalizing the criterion: a sentence is analytic if the native speaker learns to assent to it by learning one or more of its words. This accounts for such paradigms of analyticity as 'No bachelor is married', and also for the analyticity of many elementary logical truths. (...) But the definition gives no clue to the demarcation between analytic and synthetic sentences that has exercised philosophers, out beyond where anyone either remembers or cares how he learned the pertinent words. And it gives no clue, certainly, to a general concept of cognitive equivalence. (Quine, 1992: 55)

In a word, externalised analyticity has to be relativised to a speaker and to his learning of language. Granted, it is possible to 'socialise' this notion of externalised analyticity: a sentence is analytic for all members of a speaking community if it is

externally analytic for each members. But the stimulus meaning associated with each speaker is different: the history of stimulations differs in each case. In any case, this notion of meaning certainly isn't a 'realist' conception of the kind that Boghossian has in mind.

It seems to me, furthermore, that any empirical notion of meaning is bound to be unsatisfactory if intended to form the basis of an appropriate model of a priori knowledge of the sort intended by Boghossian. This is readily seen when we look at his own proposed model for our a priori entitlement to know logical laws. Taking 'C' as standing for an arbitrary logical constant and 'A' as its associated introduction and elimination rules, Boghossian (p. 357) suggests the following model:

1. If C means what it does, then A has to be valid, for C means whatever logical object in fact makes A valid.

2. C means what it does

3. A is valid

It is clear that this model relies on a strong realism about meaning and unclear how it could be otherwise.

Quine's attack on analyticity, therefore, does not amount to a complete rejection of meaning, contrary to what Boghossian's *reductio* implies. Quine's argument does go through a criticism of the notion of meaning, understood strictly in its Platonistic sense, and this does not amount to a complete rejection of meaning. To put his point differently, he is arguing that those, like Carnap, who attempt to provide a foundation for a priori knowledge by using the notion of analyticity must do so on the back of a Platonistic conception of meaning. He further maintains that if analyticity is explained in terms of that conception of meaning, it is unintelligible since there is no prospect of a clear individuation of meanings (as we have seen in connection with the identity principle). Therefore, he concludes, Carnap's conception of the a

priori collapses. Viewed in that light, Quine's attack on analyticity does not and need not involve a complete rejection of meaning. To be sure, he would accept the following version of Boghossian's argument, taking 'FM*' to mean 'facts about meaning understood Platonistically':

1. FM*1 → FM*2   Intuitive Premise

2. ~FM*2 → ~FM*1   'no entity without identity'

3. ~FM*2   No identity

4. ~FM*1   Therefore, no entity

But this clearly does not amount to a rejection of meaning as such.

To conclude, then, Boghossian's attempted *reductio* against Quine's argument ultimately fails. Quine's argument does entail the unintelligibility of a relational conception of meaning, but it does not, crucially, entail the unintelligibility of a monadic conception, contrary to what Boghossian holds. There is nothing intrinsically wrong with monadic construals for Quine since they are basically attempts to reduce the troublesome notion of meaning to something else. Now it is possible, of course, that the reducing notion turns out to be as unintelligible as the reduced notion. But it needn't be, whereas the central problem for any relational account is to devise a satisfactory criterion of individuation for one of the terms of the relation, namely the meaning M supposedly attached to an expression E.

1.6   Conclusion

The last three sections have been devoted to study and appraise counter-arguments directed to Quine's Radical Irrealism about (logical) necessity. In section 1.3, we saw that Grice and Strawson's claim that Quine's demands on explanation are unreasonable rests on a slight but pivotal misinterpretation of his views. Suitably interpreted,

Quine is best seen as trying to dissolve a picture of meaning, the Platonist picture according to which meanings are objects. In that light, asking for a suitable criterion of individuation for such entities appears to be a perfectly reasonable demand on explanation. How can we be taken to understand what a particular object is if we can't identify that very object? This consideration is what allows Quine to maintain that meanings are obscure entities. The implicit idea in play here, namely that unintelligibility follows from the lack of a clear explanation, is challenged by Grice and Strawson and examined in section 1.4. They argue that even if we lack a clear explanation for a given notion, it doesn't automatically follow that the notion is unintelligible. I suggested that a Quinean could accept Grice and Strawson's criterion of intelligibility (non-collusive assent), but all that follows from the satisfaction of their criterion in the case of analyticity is that the speakers have *a* notion of analyticity, not necessarily the notion of analyticity that Quine is attacking. Grice and Strawson need to show that the notion of analyticity prompting non-collusive assent is relevant to Quine's consideration. Boghossian follows another line of thought and argues that Quine's argument for Radical Irrealism leads to an absurdity, i.e., meaning-scepticism. Again, I have defended Quine on the basis of considerations similar to those of section 1.3.

The main upshot of this Chapter can be summed up in two points. The first concerns the dialectic of Quine's attack. We can conclude from our study of the first part of 'Two Dogmas' that his charge on analyticity goes through a criticism of the Platonistic conception of meaning, and not through a rejection of the very idea of meaningfulness. This point can easily be overlooked, partly due, no doubt, to Quine himself. The second point concerns the main target of his argument. The attack on analyticity is meant, in 'Two Dogmas', to be against the a priori / a posteriori distinction. Carnap and others attempted to ground a priori knowledge on analyticity. Therefore, if what Quine is after is a rejection of the a priori, his criticism must go through a rejection of analyticity, in so far as it construed in such a way as to ground

a priori knowledge. This is exactly what he does, and in order to do so all he needs to show is that the Platonistic conception of meaning presupposes an unintelligible notion of meaning.

His argument, properly interpreted, is not directed against the notion of meaning as such. As I mentioned before, the key for a correct appreciation of Quine's argument is to resist the temptation to inflate it. This is the mistake made by Grice & Strawson as well as by Boghossian. They took his argument to be, or to imply, the rejection of the notion of meaningfulness. They did not stress sufficiently that Quine wanted to shake the foundations of logical positivism. The result of overlooking these considerations leads one almost inevitably to inflate his argument.

CHAPTER 2

ERROR-THEORETIC IRREALISM: EPISTEMOLOGY

2.1 Introduction

As explained in the Introduction, Irrealism comes in two versions: Radical Irrealism and Error-theoretic Irrealism. Quine's argument in the first four sections of 'Two Dogmas' has often been interpreted as having to do with the *intelligibility* of a cluster of notions, and that is the reason why I called it Radical Irrealism. It is in effect a very bold claim to hold that a *prima facie* intelligible notion is in fact meaningless. We have seen in Chapter 1 that it is not clear whether Quine really advocates Radical Irrealism in 'Two Dogmas'. We saw that properly interpreted, and notwithstanding some of Quine's remarks to the contrary, the argument's target is not the notion of necessity qua notion, but rather a particular understanding of it — namely, its Platonistic interpretation. Since I find Quine's argument, so construed, compelling, it follows that I hold that a correct account of logical necessity must avoid Platonism. We have thus identified the first constraint for a correct account of logical necessity. This is a good example of the kind of positive result that can be achieved through a critical study of scepticism.

An Error-theoretic stance, by contrast, acknowledges that the concepts under discussion are meaningful. A sentence containing, say, the modal operator of necessity does make sense, that is, it makes a meaningful attempt to represent the world. The error-theorist will not challenge this. However, she holds that sentences containing modal operators are systematically false since nothing in the world instantiates the alethic modalities. Tennant offers the following characterisation of error-theory:

> (e)rror theoretic irrealism accepts the superficial syntactic appearances of our declarative sentences at face value, allows that our primitive assertions aim at the truth, but maintains that these assertions always miss their target. They are false. We can make sense of the way they represent the world as being; it is just that they misrepresent the world. They make an intelligible attempt to tell

it like it is, but they get it wrong. They picture the world as being thus-and-so,

but the world is not that way at all. (Tennant, 1997: 68).

An error-theorist about logical necessity, for instance, will hold that ascriptions of necessity are meaningful and truth-apt but that there simply is nothing which is logically necessary. Quine, it appears, can be interpreted as an error-theorist in that sense. What he does in the later sections of 'Two Dogmas', basically, is to offer a theory of knowledge doing without a specific modal notion, namely logical necessity. He does not, as is the case with other error-theorists, attempt to explain why the world does not instantiate that modal property.[1] He goes straight to offer a picture of the world and of our cognitive relation to it that has no need for logical necessity. At this point, someone could suggest that Quine is not really an error-theorist about logical necessity. The argument runs as follows. It would be consistent with Quine's philosophy that there are logical necessities. It's just that there is an account of the world doing without them. In other words, the argument goes, Quine does not have an answer to the question 'are there any logical necessities?' What he has are theories that do not, he claims, need them.

However, I think it is clear that Quine is a full-blown error-theorist in the sense that he does not believe in the appropriateness of logical necessities. Showing the superfluousness of a modal notion amounts, for Quine, to an ontological argument showing that nothing instantiates that modal property since ontological questions are settled by a theory's ontological commitments. Given that he offers a theory purged of logical necessities and assuming this theory to be elegant, explanatory, etc, it follows for him that there is no reason to believe in the existence of logical necessities.

---

[1] For instance, Blackburn in his 'Morals and Modals' (1986), who advocates a so-called 'expressivist' theory of logical necessity, which is also a manifestation of error-theoretic irrealism since it starts from the view that nothing in the world instantiates modal properties such as necessity. But it differs from Quine's approach in that it tries to explain and to justify our use of modal notions in such a way as to make it compatible with error-theory — Quine in contrast suggests that we should abandon the use of alethic modalities altogether.

Conversely, if it shown that an appropriate theory has to appeal to logical necessity, room has to be made for an account of alethic modalities and their place in our conceptual scheme.

There are two main ideas in Quine's philosophy eliciting Error-theoretic Irrealism. Firstly, there is his epistemological theory. Holism, he holds, has no need for a realm of a priori knowledge and as a result there is no role in this picture for modal notions such as necessity. Secondly, there is Quine's philosophy of logic, in particular his account of logical truth and logical consequence. He advocates a *substitutional* account of logical truth in which, he maintains, no use of intensional notions such as necessity or analyticity are in play. Both theories are manifestations of Error-theoretic Irrealism about logical necessity. It is the aim of this Chapter as well as the following two to show that Quine is wrong in holding that holism and the substitutional account of logical truth are free of logical necessity. More precisely, I want to defend the idea that although there are genuine insights associated with holism and Quine's philosophy of logic, it is a mistake to hold that these important and illuminating philosophical ideas give rise to Error-theoretic Irrealism.

We'll proceed as follows. The next section studies the relation between the two dogmas and the link between them and holism. Section 2.3 is devoted to expose Quine's holistic picture of knowledge as sketched in sections 5 and 6 of 'Two Dogmas'. Section 2.4 attempts to locate logical laws in Quine's holism, and identifies a problem with this. The next section addresses this problem via Lewis Carroll's regress. It then becomes clear that the problem outreaches the bound of Quine's philosophy and raises a pressing question that has to be answered no matter what your philosophy of logic is. Accordingly, various solutions and lines of research are studied. In the end, I settle on a Davidsonian solution. In the last section, I examine Shapiro's claim that Quine's epistemological holism requires that some statements are fundamentally a priori in the web of belief, in particular the question whether a Quinean could accept

that a minimal set of principles be understood as being a priori. This Chapter sets the ground for the next, in which Wright's argument against Quine's holism is analysed.

2.2 The two dogmas and holism

In this section, I want to clear the ground and clarify some important ideas by contrasting them with one another. Here are the pairs that will be put under scrutiny:

- The relation between the two dogmas.

- The relation between holism and the two dogmas.

- The relation between the first dogma and Error-theoretic Irrealism.

- The relation between holism and logical necessity.

Disentangling these ideas proves to be helpful in understanding how, in particular, the abandonment of the two dogmas relates to Quine's philosophy of logic. Firstly, though, a word on the second dogma is in order. In section 5 of 'Two Dogmas', Quine attacks the idea that the meaning of a statement is determined by specifying its method of verification. This theory is in effect a reductionist theory of meaning, reducing the meaning of an expression to either an expression about immediate experience (radical reductionism) or to an expression about confirmation or refutation (attenuated reductionism). Quine offers an argument to refute radical reductionism; but concerning the attenuated form of reductionism he offers instead a theory (holism) which does not entail it. Now that we have a 'working idea' of the second dogma, we are ready to examine the conceptual relations between the notions mentioned previously.

2.2.1 The relation between the two dogmas

The two dogmas attacked by Quine in his paper are, respectively, the analytic / synthetic distinction (first four sections) and the verificationist theory of meaning (fifth section). How should we understand the relation between the two dogmas? Quine says that the two dogmas are "at root identical" (1951: 41). What does he mean by that? They appear, on the face of it, to be independent. But he goes on to explain that if a statement acquires meaning by relation to its method of verification (second dogma), then it is a small step to regard some statements as verified no matter what — and these are the analytic / necessary statements (first dogma) (1951: 41). The reasoning here is, presumably, that to hold that it makes sense to speak of all meaningful statements being somehow reducible in principle to statements which are empirically verifiable renders plausible the idea of a class of statements which are always verified — the necessary statements. Notice moreover, as pointed out in Chapter 1, that Quine clearly conflates the analytic / synthetic distinction with the necessary / contingent one — for him these dogmas are in effect one and the same dogma. Since he also argues, as we have just seen, that the dogma of the analytic / synthetic distinction is identical to the dogma of verificationism, we have it by association that for Quine, the dogma of the necessary / contingent distinction is identical to the dogma of verificationism.

But there are problems with these associations. *Firstly*, verificationism is compatible with the view that there are *no* statements verified no matter what — no necessary statements in Quine's sense. That is, it is not clear that espousing verificationism commits one to believe in a set of necessary statements so that the dogma of verificationism entails the dogma of the necessary / contingent distinction. Indeed, it could be the case, for all we know, that all meaningful statements must be capable of verification to be meaningful, but that there are no statements somehow always verified. On the face of it, verificationism is compatible with the view that there are

no 'necessary' statements, only contingent ones. What verificationism allows is the mere conceptual possibility of necessary statements, not their existence. Again, if we keep in mind that Quine's target is Carnap, the assumption that verificationism and acceptance of necessary statements go hand in hand is understandable — but it is an assumption. It requires a further argument to show that verificationism entails the existence of such statements.

Secondly, one can easily grant the possibility of statements verified no matter what, but resist the claim that such statements are analytic. It is not clear whether acceptance of the existence of statements verified no matter what entails a commitment to the analytic / synthetic distinction. The wanted conceptual connection here between 'verified no matter what' and analyticity is missing. It takes a special theory, such as logical positivism, to spell it out.[2] For instance, it seems to be a coherent view that 'Water = H_2O' is necessary without being analytic. So it is far from being a truism that endorsing the necessary / contingent commits one to endorse the analytic / synthetic distinction. The issue is complicated due to the fact that Quine conflates the analytic/ synthetic distinction with the necessary / contingent one. Quine's 'small step' from reductionism to analyticity is in effect tantamount to a fully developed logical positivist philosophy.

Moreover, there is no reason to maintain that the identity between the dogma of the analytic / synthetic distinction and the dogma of verificationism holds. To accept the analytic / synthetic distinction is certainly not to accept ipso facto the verificationist theory of meaning. If that were the case, it would be disastrous for the contemporary believers in analyticity, such as, for instance, Boghossian. Virtually no one today subscribes to the verificationist theory of meaning — we can safely say that it is a rare instance of a falsified philosophical theory.

---

[2]That is, logical positivism à la Carnap associates analyticity with necessity in a theoretical framework, thus explaining and defending the connection between the two concepts.

To sum up, we have seen that Quine holds that the three dogmas are identical. But we have just seen that it is possible that:

1. the dogma of verificationism does not entail the dogma of the necessary / contingent distinction

2. the dogma of the necessary / contingent distinction does not entail the dogma of the analytic / synthetic distinction

3. the dogma of the analytic / synthetic distinction does not entail the dogma of verificationism.

The truth of any of the three above renders false the claim that the three dogmas are identical. In the end, therefore, the two dogmas (strictly speaking, the three dogmas) are clearly not identical despite their close relationship.

2.2.2 The relation between holism and the two dogmas

It is in the fifth and sixth sections of 'Two Dogmas' that Quine puts forward a picture of our cognitive relation to the world allegedly purged of the notions of analyticity, necessity, and aprioriticity. These sections articulate a view of knowledge — holism — in which use of these intensional notions is supposedly superfluous.[3] Quine thus holds a holistic theory of knowledge that does without the dogmas. However, let us not forget that holism does not imply, taken by itself and independently of the rest of Quine's paper, the unintelligibility of intensional notions. It is in effect clearly possible to advocate holism and to acknowledge the intelligibility of analyticity and necessity. The holism would be a local holism, with some statements being held true come what may (most likely statements of logical consequence and some basic logical

---

[3]I shall have more to say on holism in the next section. For the moment, however, let's notice that the focus in this Chapter will be on *epistemological* holism rather than *meaning* holism.

laws). Quine, of course, advocates a full-blown holism. But it is important to keep in mind that holism can come in degrees.

The opening sections of Quine's paper are mainly concerned to show the unintelligibility of (an understanding of) the notions of analyticity and necessity, whereas his paper as a whole consists essentially of two arguments: criticism of the two dogmas and defence of holism. How are we to understand the relation between the two main arguments? It is certainly not a mere coincidence that holism is advocated right after the criticism of the two dogmas. There are two options to consider:

1. Holism 'follows' from the abandonment of the two dogmas.

2. The two main arguments are independent.

Interpretation (1) makes too strong a connection between the two arguments. We have to keep in mind that Quine is attacking a version of Platonism for necessity, meaning, and analyticity (first dogma) and an untenable view about meaning (the verificationist conception of meaning). It is hard to see how holism could follow from the abandonment of these two highly controversial theories. There is plenty of theorising space for a *non*-holistic view of knowledge that does without both Platonism and the verificationist theory of meaning. At the very least, it would require an additional argument to show that holism is the logical outcome of the critical parts of 'Two Dogmas'. Interpretation (2) thus has more plausibility. However, to hold that the two main arguments of 'Two Dogmas' are logically independent need not suggest that they are unrelated. Quine's main target in sections 1-4 of his paper is, as we have seen, a priori knowledge. He claims that the notion of meaning inherent in, for instance, Carnap's conception of the a priori can't do its explanatory job. Instead of trying to redefine a priori knowledge on the basis of concepts other than meaning and analyticity — as he could conceivably have done — Quine puts forward a view of our cognitive relation to the world (holism) which does not make appeal to this kind of

knowledge. So the two arguments *are* related: the critique of a priori knowledge can be seen as somehow provoking holism without, strictly speaking, entailing it.

If the two arguments are independent in the sense just explained, it follows that Error-theoretic Irrealism and Radical Irrealism are also logically independent. On this understanding — the one I will adopt here — Error-theoretic Irrealism does not, by itself, imply any thesis about the intelligibility of logical necessity. The alternative reading (1) would be possible only if Quine was arguing, in sections 1-4 of 'Two Dogmas', that the notions of analyticity, necessity, and meaning are by themselves unintelligible. If he had shown this, then it could be said that in a sense, holism follows from the critique of analyticity. Let me explain. The correctness of an attack on the very intelligibility of analyticity would show the deep need for a picture of knowledge doing without this notion. It could be the case that a picture of knowledge purged of analyticity other than holism is appropriate — but at the very least Quine's criticism of the two dogmas, so interpreted, would directly prompt a picture of knowledge entailing Error-theoretic Irrealism. We have seen however that this is not what Quine does. His main concern in the first four sections of 'Two Dogmas' is to criticise a particular *interpretation* of this family of concepts, not their very intelligibility. As such, this criticism leaves conceptual space for an epistemological theory using another interpretation of the concepts involved in the criticism (meaning, analyticity, necessity). This explains why, in the end, the two main parts of Quine's classic paper are independent. The key point is to remember that Quine's attack precludes only a picture of knowledge which makes crucial use of the notion of the a priori cashed out in terms of meaning and analyticity. Whether Quine would accept an alternative view using a differently constructed notion of a priori knowledge is another story, which will be discussed in section 2.6.

2.2.3   The relation between the first dogma and Error-theoretic
Irrealism

In a key passage, Quine tells us that if we drop the analytic / synthetic distinction, there will be no principled distinction between contingent statements and those that are true 'come what may' (1951: 43). This passage appears to contradict the claim I made in the previous paragraph, in that it states the inevitability of a picture of knowledge doing without necessity. Two comments on this. First, Quine is here making an inferential step which he does not justify — he has not shown to us that the analytic / synthetic distinction as such is meaningless, only that a particular understanding of it is. It can hardly be said — without further explanations — that to abandon the Platonistic construal of analyticity entails the annihilation of the necessary / contingent distinction. Second, it is controversial whether the abandonment of the analytic / synthetic distinction as such (that is, independently of its construal) leads directly to the abolition of a principled distinction between contingent and necessary statements. It is often said, for instance, that the fact that water is composed of H_2O is a necessary one, but 'water is H_2O' is certainly not an analytic statement — rather it is an empirical 'fact' that water is constituted of this particular combination of molecules.[4] It is therefore a possibility to drop the analytic / synthetic distinction without abandoning the necessary / contingent distinction. Despite what Quine says, Error-theoretic Irrealism of the kind under investigation in this thesis clearly is not, in a strict sense, a consequence of the critique of analyticity.

2.2.4   The relation between holism and logical necessity

Now that we have explained the relation between Error-theory (holism being a version of it) and the two dogmas, we now turn to another important relation, namely

---

[4]See for instance Kripke (1971: 122-123, footnote 63) and Putnam (1983) for a discussion of the relation between necessity and analyticity along these lines.

the relation between holism and logical necessity. Although we will discuss the content of Quine's holism in the next section, we shall now be interested in the conceptual relation between it and logical necessity. We know that holism supposedly 'does without logical necessity' — this is why it is a form of scepticism. In exactly which sense, however, is this locution to be understood? That is, in which sense does holism imply Error-theoretic Irrealism about logical necessity? This is a different question from the previous one: we just saw how holism and *Radical* Irrealism are independent, and now we're asking what connects holism to *Error-theoretic* Irrealism. Holism — in its extreme, radical form as exposed in section 6 of 'Two Dogmas' — is compatible with both Radical and Error-theoretic Irrealism. Quine himself is an epistemological holist who advocates a brand of Radical Irrealism and Error-theory about logical necessity. To be more precise, the radical form of holism implies Error-theory, but does not imply Radical Irrealism, even though they are both compatible with holism.

Quine begins with the assumption that there is no principled distinction between contingent statements and necessary ones (we just saw that he has not shown this to be the case, but this is another story). This starting point is in effect a manifestation of Error-theoretic Irrealism, and holism is Quine's way of explaining knowledge whilst espousing Error-theoretic Irrealism. It is important to remember that Quine starts with a strong prejudice against intensional notions. For him, the more we can do without them, the better.[5] In a sense which will be explained in the next section, holism, Quine holds, has no need for the notion of necessity. Holism is therefore a natural, if not logical, product of the abandonment of the contingent / necessary distinction — independently of how we are to understand these notions.

In effect, if there is to be no principled distinction between statements such as

---

[5]Quine's dislike of intensional notions is scattered throughout his writings. See in particular (1960: Chapter 6 ('Flight from Intension'); 1995a: Chapter 8). To get a sense of how important the principle of extensionality is for Quine, see (1953c: 160-161). As early as in his (1941: 16-17), he expressed criticisms of modal notions and of breach of extensionality.

'the cat is on the mat' and statements representing a logical derivation, it follows that there is no justification for asserting statements such as $\square(\Gamma \vdash_L \Theta)$. Even a logical derivation — a proof — cannot be ascribed necessity since nothing distinguishes it from a contingent statement. It is at this point that holism becomes an attractive option. There might be alternative pictures of knowledge entailing Error-theoretic Irrealism. But holism clearly is a natural candidate. Quine needs a picture of knowledge in which there is no role for the notions of necessity and analyticity, and holism provides just that, together with an empiricist flavour. I think the way to understand why holism is so congenial to Quine is a mixture of these two ingredients: repudiation of the existence between necessary and contingent statements together with a taste for empiricism.

2.3 Quine's holistic picture of knowledge

Now that we have a clear view of the place of holism in Quine's philosophy and its relation both to Error-theoretic Irrealism and to the criticism of the two dogmas, it is time to examine in what exactly holism amounts to. In 'Two Dogmas' Quine writes:

> The totality of our so-called knowledge or beliefs, from the most casual matters
> of geography and history to the profoundest laws of atomic physics or even of
> pure mathematics and logic, is a man-made fabric which impinges on experience
> only along the edges. (Quine, 1951: 42)

The first impression given by this passage is that Quine is here talking about epistemological holism, in contrast to meaning holism. The distinction between these two kinds of holism is not clear-cut; indeed, they seem very closely related. In a nutshell, however, meaning holism is a view explaining how words and sentences get their meaning. Meaning holism is the view that understanding a sentence requires

understanding a chunk of language. How large this 'chunk' is will be variable. For some holist philosophers, the understanding of a concept requires the understanding of the whole language. For others, it requires only understanding of the chunk of language related to the concepts involved. By contrast, epistemological holism is a view about how our language relates to the world. It is also an attempt to explain how words get their meaning, but the emphasis this time is on the way language interacts with the world and how the world has an impact upon the meaning assigned to the expressions of the language. To put it simply, meaning holism is a view about how words get their meaning; epistemological holism goes beyond that and asks what is the relation between language and the world. The answer provided by epistemological holism is that our language faces the world as a corporate body. When one is talking about meaning holism, one is more concerned about the relation between expressions than in the relation between language as a whole and the world. In 'Two Dogmas', as we shall see, the focus is on epistemological holism. In what follows, this is what we shall understand by 'holism'.

Let us now return to the quotation above. It reveals two important features of holism: one, that knowledge is the product of mankind, and second, that holism à la Quine is of an empiricist sort. I think the first feature is to be taken in a strong sense: Quine wants to convey the idea that what is true has nothing to do with 'correspondence', but rather solely with explanatory value. This is what makes Quine a pragmatist, although he espouses a very different pragmatism than the one advocated by the logical positivists. He is also, of course, an empiricist, to the extent that for him, *all* statements bear a relation to experience. It might be illuminating to contrast Quine with Carnap at this point. Carnap is also an empiricist to the extent that he considers as being meaningful exclusively statements which are somehow reducible in principle to directly verifiable statements. But statements about logic and mathematics, not verifiable according to Carnap, are necessary and analytic,

outside the realm of the empirical. Quine does not make such a distinction — his empiricism, as well as his pragmatism, differs from Carnap in important respects. For Quine, contrary to Carnap, there are no statements 'outside the realm of experience'. Admittedly, logic and mathematics are not along the edges of the man-made fabric, so their relation to experience is remote. But nevertheless they are, in the end, hostage to experience in the sense that for Quine, the web of belief is an instrument used to predict experience — or, to talk more precisely, to predict observation sentences. Logic and mathematics owe their 'truth' to their usefulness in predicting observation sentences and organising them into a coherent whole. We thus see how Quine's pragmatism and empiricism are closely related.

It is also important to bear in mind that the point just made, namely that all statements bear a relation to experience for Quine, is not to be confused with the idea that each individual statement has an individual, isolated empirical content. To say this would be to recoil into the verificationist theory of meaning. One of the key theses of holism is precisely that "(t)he unit of empirical significance is the whole of science" (Quine, 1951: 42) and that "our statements about the external world face the tribunal of sense experience not individually but only as a corporate body" (Quine, 1951: 41). This is a straightforward consequence of the abandonment of the second dogma. If one is to maintain empiricism whilst denying that each individual statement has an empirical content, the natural option is to hold that we must rather speak of 'chunks' of statements having an empirical content: that seems the only way to retain the empiricist insight that experience has a crucial impact on the content of our beliefs without embracing the verificationist theory of meaning. To come back to the relation between holism and the two dogmas examined in the previous section, it appears that a particular aspect of holism — that statements face experience *en bloc* — is somehow entailed by the rejection of the second dogma if we add empiricism as a premise.

But how, precisely, does experience 'impinge' on the set of our beliefs? It becomes useful at this point to evoke the picture of the web of belief, even if this particular metaphor does not surface in 'Two Dogmas'. In this picture, statements are nodes in the web, and they are connected together via their logical interconnections. Statements situated near the periphery are more directly related to experience; statements near the centre of the web are less related to experience. A statement such as 'this shirt is blue', for instance, will be peripheral compared to '2 + 2 = 4'. The fact that 'this shirt is blue' is situated at the periphery of the web merely means that its truth-value is more likely to be changed in the face of a recalcitrant experience than '2 + 2 = 4' is. Faced with an experience that contradicts 'this shirt is blue', we are much more likely to change the truth-value of this particular statement than statements which are connected to it but are situated at the centre of the web. This is because we tend to modify our beliefs in a conservative manner: changing the truth-value of a peripheral statement will not disturbingly affect the rest of the web, contrary to what happens if we change the truth-value of, say, arithmetical or logical statements. To sum up, the difference between an observation sentence and a logical truth is one of degree, not of kind. According to Quine, the belief in the necessary / contingent distinction is the product of elevating statements situated at the centre of the web — as empirical and contingent as the others — to a special status, thus bestowing them a privileged standing. This unwarranted elevation of such statements is to be explained by a conflation of psychological certainty with necessity: an unjustified flight towards intension.

We have to keep in mind that when faced with a recalcitrant experience, what is contradicted is not merely 'this shirt is blue' or any other individual statement, but all the statements logically connected to it — despite the fact that a recalcitrant experience sometimes seems to contradict only one statement (the observation sentence). It explains the possibility, in Quine's picture, of a change in the truth-value

of any statement, including those at the centre: it is possible, in the first instance, because these latter statements are logically interconnected with those at the periphery and, secondly, because they confront experience together with the peripheral ones as a whole. This is the important lesson to draw from the abandonment of the verificationist theory of meaning. For Quine, it doesn't make sense to talk about the empirical content of a statement in isolation from other statements.

Under the Quinean picture, what happens in the face of a recalcitrant experience is this. The recalcitrant experience contradicts a set of statements that are true according to some theory T. It is a set since statements, remember, face the tribunal of experience as a corporate body. In 'Two Dogmas', Quine writes as if it were the whole web of belief that confronts experience at any time. In later writings,[6] Quine espouses a less radical holism (called 'moderate holism') — it makes more sense to say that it is a set of sentences, rather than the whole web, that confronts a given recalcitrant experience (or any experience, for that matter). For our purposes, however, this distinction does not matter: all we need is the idea of a set of statements confronting experience. A recalcitrant experience is an experience that, by definition, does not 'fit' a particular set of sentences.[7] In order to accommodate that experience, changes must be made. That means that there ought to be a redistribution of the truth-values of some sentences in the relevant set. For example, one might decide to change the truth-value of an instance of 'this shirt is blue', or of 'this is the same object as I saw two minutes ago' in order to accommodate a particular experience conflicting with a set of sentences: let's say that this recalcitrant experience consists in one seeing what seems to be the same shirt but in a different colour. One will naturally be reluctant to abandon the law of non-contradiction, such that the shirt could be

---

[6]See for instance Quine (1995a: Chapter IV). For further discussion on moderate holism and Quine, see Hookway (1988: Chapter 2).

[7]This picture of what happens when facing a recalcitrant experience will be refined and discussed at length in the next Chapter.

said to be both blue and not blue and thus 'save the phenomenon'. Abandoning the law of non-contradiction would in effect occasion a radical re-evaluation of truth-value assignments. According to Quine, it is a matter of conservativeness and simplicity that guides us through the redistribution. We are more likely to change the truth-value of observation sentences than to change the truth-value of the law of non-contradiction simply because the former alternative occasions less disturbance inside the web.

Quine's pragmatism is thus of a more extreme kind than Carnap's: not only is the choice of a conceptual scheme (a logic) a matter of choice guided by pragmatical considerations such as simplicity and conservativeness (as in Carnap), but so is, ultimately, the truth-value of all statements inside the web — including observation sentences. Carnap, for instance, holds that the choice of a logic is a matter of choosing a particular conceptual scheme. For Carnap, choosing a conceptual scheme is an external question, and as such it is not an empirical question.[8] This explains why, for Carnap, the reasons for adopting a particular conceptual scheme are pragmatic reasons, not empirical ones. Once a conceptual scheme is chosen, experience has no import whatsoever on the truth or falsity of logical and mathematical statements since they are on this picture analytic and necessary. Internal questions, however, are empirical and there is simply no pragmatism involved in deciding their truth or falsity. The choice of a conceptual scheme brings with it a set of necessary and analytic truths, and the problem of identifying the necessary statements is reduced to the problem of reading off the conceptual scheme and identifying which statements are necessary in that conceptual scheme. In Quine's system, on the contrary, there simply is no place for necessary statements — all statements relate to experience and

---

[8]See Carnap (1957) on this particular point. For a good discussion of Carnap's internal / external distinction with regard to scepticism, see Stroud (1984, Chapters 4 and 5). For Carnap, the question 'are there logically necessary statements?' is meaningless qua philosophical question, that is, qua external question.

it is pragmatic considerations that guide us not only in the choice of a conceptual scheme but in the choice of the statements all over the web. As Hookway (1988: 50) put it: "The closing sections of both 'Two dogmas' and 'On what there is' are full endorsements of pragmatism: Carnap's pragmatic considerations become relevant to all questions once the contrast between external and internal is abandoned". [9]

This is a strong, compelling metaphor for how knowledge is acquired and modified, one that has of course proved influential. The crucial question for us is: what is the status of statements of validity in this picture? That is, what is the status of a statement representing a logical proof in the form $(\Gamma \vdash_L \Theta)$ as explained in the Introduction? A non-sceptic wants to say that such statements are logically necessary. But Quine holds, as we have seen above, that there are no true logically necessary statements. However, there are for Quine true *simpliciter* statements of the form $(\Gamma \vdash_L \Theta)$, statements to which a non-sceptic would typically be inclined to ascribe necessity — to assert their validity in a strong sense. That is, according Quine there is no true statement of the form $\Box(\Gamma \vdash_L \Theta)$, but there are true, 'unboxed' statements of the form $(\Gamma \vdash_L \Theta)$. For instance, even for Quine it is true that if one starts with such and such axioms and rules of inferences, such and such theorem follows. What are the place and the role of such statements in the web of belief? We shall be concerned with this question in the remaining of this Chapter.[10] Let us analyse this crucial passage to begin with:

---

[9]One can view the situation historically like this. Carnap denied the Kantian idea of *synthetic* a priori knowledge: for Carnap, there is only *analytic* a priori knowledge. Kant was trying to delineate a broad field for a priori theorising, thus vindicating a whole area of philosophical research. Carnap restricted the sphere of philosophy to analytic propositions. Quine goes even further and eliminates the frontier of a priori knowledge by deconstructing the very concept of an analytic / synthetic distinction. By doing so, Quine can be seen in a sense as leaving no area for specifically philosophical theorising. This is coherent with his idea of a naturalized epistemology, in which philosophy is but a branch of cognitive science. The present investigation can be regarded as a small but crucial step in claiming back the specificity of philosophical research.

[10]The next Chapter analyses a related argument by Wright showing that such statements must have a special status for one who endorses holism, whereas here we are interested to locate these statements in the web.

> Reëvaluation of some statements entails reëvaluation of others, because of their logical interconnections — the logical laws being in turn simply certain further statements of the system, certain further elements in the field. Having reëvaluated one statement we must reëvaluate some others, which may be statements logically connected with the first or may be the statements of logical connections themselves. (Quine, 1951: 42)

We can almost see here Quine's struggle to come to term with the status of the logical connections in the web of belief. We can be certain of one thing: for Quine, the *logical laws* are statements in the web, and as such they have the same status as, say, observation sentences. The only feature distinguishing logical laws from observation sentences is their position in the web: the former are located in the centre, and the latter are located at the periphery. This, at least, is the most plausible interpretation of what Quine is trying to explain in the quoted excerpt.

However, upon close examination, things are not so clear-cut. It would indeed be straightforward to analyse the components of the web according to the metaphor of 'distance from the periphery', with statements typically close to the periphery such as observation sentences and others, such as logical laws, situated at the centre. The web would be a seamless whole, containing thousands of sentences, all of which are logically interconnected. I wish to argue that this web of belief picture has no place for logic, contrary to what Quine holds. In particular, Quine's claim that logical laws are merely further statements in the web cannot be sustained. In order to see this, however, we will have to take a detour to Quine's logical system and examine the notation he uses in his system as well as the status of the logical laws in this system.

2.4 The web and its logical laws

Let us reflect, to begin with, on the following. What does Quine have in mind
when he talks about 'logical laws'? What is a logical law for Quine? When we think
of a logical law, we have in mind the laws of excluded middle, non-contradiction,
DeMorgan's law, and, also, modus ponens (although this one is a special case, as we
will see later). Logical laws are general: it is in the nature of logic that a logical law
is uninterpreted.[11] That is to say, logic is about everything conceivable — it is about
the relations holding between statements of any kind. In other words: there is no
intended interpretation for logic. The law of excluded middle, for instance, applies
to all statements. More exactly, we want to say that the fact that every statement
in the web is either true or false is due to the principle of bivalence via the law of
excluded middle. The principle of bivalence, of course, is understood today as an in-
terpretation of the law of excluded middle, an interpretation based on truth-functional
semantics.[12] Quine advocates truth-functional semantics, so for him acceptance of
excluded middle is acceptance of the principle of bivalence. Quine himself, however,
does not distinguish between the law of excluded middle and bivalence (Quine, 1970:
83).

Now let's try to locate logical laws in Quine's holism. We will focus here on the
law of excluded middle for expository purposes. What would the statement of ex-
cluded middle be like? For Quine himself, it is illustrated thus:

(**EM**) $\lceil \phi \text{ v} \sim \phi \rceil$

This way of illustrating excluded middle uses the notational conventions of *Mathe-*

---

[11]On the generality of logic, see Quine (1995a: 52).

[12]For the distinction between logical laws and their corresponding semantic principles, see Dum-
mett (1978: xix).

matical Logic, which I take to be the clearest and most systematic notation by which to formalise logic.[13] The corner quotes are meant to clarify the fact that we don't want to say that it is the statement referred to by quoting '$\phi \vee \sim \phi$' which is true. In effect, '$\phi \vee \sim \phi$' is merely another statement, conveying no generality whatsoever. It just says that either the inscription 'ϕ' is the case or the inscription '$\sim \phi$' is the case — it is meaningless to assert '$\phi \vee \sim \phi$' if we stick to the usual way of using quotes. The Greek letters, in Quine's notation, serve as names of unspecified statements. Quasi-quotation is meant to convey the generality intended when we talk about excluded middle. (**EM**) refers to any statement obtained by substituting uniformingly the occurrences of the propositional variable 'ϕ' in $\lceil \phi \vee \sim \phi \rceil$ with sentences.

It is important to recognise that the generality exemplified by the use of the corner quotes is crucial here. This generality is intended to capture the insight that logic is an uninterpreted theory. More importantly, in connection to holism and the web of belief, the generality reminds us that logical laws apply to every statement in the web. This, as we're about to see, is the crux of the problem.

However, we haven't arrived yet at a proper formulation of the law of excluded middle with $\lceil \phi \vee \sim \phi \rceil$. This statement refers to any expression obtained by replacing uniformly the Greek letters by names of specific statements, such as 'the cat is on the mat'. It does not say that all such substitutions are true — this is what is needed for transforming $\lceil \phi \vee \sim \phi \rceil$ into the *law* of excluded middle. In Quine's notation, the law of excluded middle is exemplified by:

(**LEM**) $\vdash \lceil \phi \vee \sim \phi \rceil$

We get this result via the axiomatic metatheorem *100 of *Mathematical Logic*:

---

[13]The law itself appears in (1940: 51). Quine explains his notational conventions in §§5 and 6 of that same book.

(*100) If ϕ is tautologous, $\vdash \phi$.

What (**LEM**) says is that any result of admissible substitutions in $\lceil \phi \vee \sim \phi \rceil$ is a theorem (a logical truth). The turnstile, applied to (**EM**), transforms it into a logical law. (**LEM**) — and *100 — is a *meta*theorem since it governs the condition under which a statement can be a logical truth. A logical truth is composed of proper statements, in contrast to statements of the metatheory which are bracketed between corner quotes. For instance, 'the cat is on the mat or the cat is not on the mat' is a logical truth, but '$\vdash \lceil \phi \vee \sim \phi \rceil$' is a metatheorem. The statement 'the cat is on the mat or the cat is not on the mat' is logically true in virtue of (**LEM**). (**LEM**) is therefore a law regulating all statements in the web, policing, as it were, all statements from a privileged, external standpoint. This is ultimately the source of the problem.

Let us recapitulate. In Quine's web of belief, there are many kinds of statements. There are for instance:

1. observation sentences — 'the cat is on the mat'

2. logical truths (theorems) — 'the cat is on the mat or the cat is not on the mat'

3. tautologous forms — $\lceil \phi \vee \sim \phi \rceil$

4. logical laws — $\vdash \lceil \phi \vee \sim \phi \rceil$

5. axiomatic sentences —If ϕ is tautologous, $\vdash \phi$.

Statements such as (4) and (5) settle the logical truths (the theorems) in the web of belief. That is to say, the metatheorems state the conditions for being a theorem. Therefore, in order to be a logical truth, a statement must satisfy the conditions imposed by the metatheorems. It is the metatheory, presumably, which determines

the logical connections in the web of belief. So there ought to be, it appears, levels in the web: it can't be a seamless whole. Some statements (the metatheorems) seem to occupy a special place.

This patent need for a metatheory raises many problems. How does the metatheory relate to the web? What is the relation between the metatheory and the rest of the web? Does it make sense to hold that all statements are logically interconnected in the web, including the metatheory? The metatheory is *about* the 'rest' of the web — in what way are they logically connected? It looks as if the web has to be two-dimensional. On the one hand we have the web as we know it, with its observation sentences, its scientific theories, its logical truths. One the other we have the metatheory regulating the logical interconnections between these sentences. The next two sections investigate these thoughts.

2.5   Quine and the Tortoise

The difficulty we have to address is to assess how the statements of the metatheory relate to the other statements, more precisely how the metatheorems relate to the logical truths in Quine's system. This problem will be tackled by studying one of the most puzzling problems in the foundations of logic — a problem which, I believe, Quine's philosophy of logic is unable to solve. Lewis Carroll presents the conundrum in question in his 'What the Tortoise said to Achilles'. The strategy in the following subsections will thus be to spell out this problem as clearly as possible and then explain why Quine's philosophy of logic cannot solve it. Following that, I will survey other possible solutions to this problem and subsequently offer a way out.

2.5.1   The regress

Let us consider this simple argument:

(1) If Socrates is a man, then Socrates is mortal.

(2) Socrates is a man.

(C) Therefore, Socrates is mortal.

Here, we want to say, intuitively, that you are warranted in drawing the conclusion (C) on the basis that the premises are true and also on the basis of the logical law of modus ponens. Indeed, staying at the intuitive level, mere warrant for the premises is not enough to draw the conclusion: you must also *know* that the premises possess a valid logical form.

How should this knowledge be represented? Well, you must know that modus ponens is a valid logical form, that is, you must know:

(**W**): If an argument displays the logical form of modus ponens, then it is valid.

(**W**), then, is to be understood as the *warrant* enabling you to 'go' from the premises to the conclusion in our original argument. The question now is: how does knowledge of (**W**) warrant you in drawing the conclusion (C)? It seems that you need to know at least two things. One, that (**W**) is correct and second, that the argument above is an instance of modus ponens. So what you need to know can be represented like this:

(a) (**W**): If an argument displays the logical form of modus ponens, then it is valid.

(b) The argument above ((1), (2), (C)) is an instance of modus ponens.

(c) Therefore, the conclusion (C) — that Socrates is mortal — can be drawn.

The problem is immediate: the step from (a) and (b) to (c) is itself inferential. So in order to be warranted in drawing the conclusion (c), you need to possess an additional piece of knowledge, namely:

(i) (**W**): If an argument displays the logical form of modus ponens, then it is valid.

(ii) The argument above ((a), (b), (c)) is an instance of modus ponens.

(iii) Therefore, the conclusion (c) — that the conclusion (C) can be drawn — can be drawn.

It seems that this pattern could be repeated indefinitely; if so, it means that claiming knowledge of (**W**) when drawing inferences leads to a regress of warrants. This in turn would mean that explicit knowledge of the validity of the modus ponens schema cannot act as a warrant for the conclusion (C).

2.5.2   Internalism, externalism, scepticism

Let us think carefully about the conditions under which there is a regress. It seems that in order for the regress to go through, it must be a requirement that the inferrer possesses a warrant for drawing the conclusion. In other words, the inferrer needs to *know* that modus ponens is a valid rule of inference. This additional piece of knowledge is illustrated by (**W**). However, a further requirement is needed. It must be the case that claiming knowledge of (**W**) is always necessary — that is, it must be a reasonable demand to always require, on the part of the inferrer, knowledge of the validity of modus ponens. The first requirement amounts to what Wright labels 'simple internalism'. It is to be defined thus:[14]

Simple internalism: In order to acquire knowledge by inference, the thinker must

---

[14]The formulations of simple internalism and simple externalism are borrowed from Wright (forthcoming).

possess a warrant to the effect that the rule used is valid.

The second requirement seems plausible — given that one espouses simple internalism. Indeed, if the inferrer must possess an appropriate warrant to acquire knowledge by inference in one case, then there seems to be no reason why we should not require such a warrant in the other cases. That is, if knowledge of the validity of modus ponens is required to draw the inference (C), then it is also required to draw the inferences (c) and (iii). There does not seem to be any room for an internalist to stop the regress by saying that the warrant is only needed in the first instance but not in the others.

So it seems that the simple internalist position as outlined above is subjected to the regress. In contrast, a 'simple externalist' standpoint about inference does not face the threat of the regress. According to this externalist position, an inferrer does not need to know that modus ponens is a valid rule of inference in order to draw the conclusion (C): all that is required is that the rule is in fact valid — knowledge is not a necessary condition. Simple externalism is characterised thus:

Simple externalism: In order to acquire knowledge by inference, the rule used must be valid. Knowledge of this rule is not necessary.

It is clear that one who espouses this view will not be bothered by the regress, since it is not necessary to be capable of producing a warrant such as (**W**) in order to draw the inference. All that is necessary is that the rule is indeed valid. In this sense, the challenge posed by Carroll's regress is one that will affect only a broadly internalist conception of inferential knowledge.

There is another stance about our inferential practice that is not subjected to the regress — namely, scepticism about warrant. According to this view, when one makes an inference, there is no warrant in play. That is, on this view, inferring the way we do is not governed by any norm. Talk of a rule being 'valid' is nonsense according to the sceptic. We will come back to scepticism in later subsections, but let us define it:

Scepticism: Inference is not guided by any kind of warrant (internalist or externalist) and there are no norms of correct reasoning.

This is a sceptical view about our normal conception of inferential practice in the sense that it purports to show the superfluousness of there being a norm that we follow when we draw inferences. Indeed, talk of warrant easily transfers into talk about norms. The warrant — say, (**W**) — allows me to draw the inference, and that warrant is a *norm* that I follow. So the sceptical predicament is that inferential practice is not normative, if we understand norms in the way just suggested. That, it seems to me, is a highly unattractive take on the matter.

2.5.3   Setting up the dialectic

Carroll's regress seems to establish that knowledge of logic (that is, in this case, modus ponens) is unable to warrant simple inferences — but this is, arguably, precisely the role of logic. Invoking one of the most fundamental laws of logic is of no help to warrant the logical move from the premises to the conclusion. So according to Carroll's story so understood, it is impossible to be internally warranted in drawing inferences. There is an infinite regress of necessary warrants. That is a point deserving emphasis: what Carroll's regress purports to show is not, as it is often assumed, that the mind can't move — the mind *does* move, we have no problem in drawing simple inferences. It's rather about the *ground* of the move, about whether (or how) logic can *make* the mind move. We usually assume — in this kind of case — that our move from the premises to the conclusion is warranted by appealing to logical laws. But Carroll's example purports to show that this a misunderstanding of our inferential practice. The move, if we accept the regress and its conclusion, is *not* warranted the way that the simple internalist model requires .

We can view the dialectic in regard to Carroll's regress like this. Prima facie, we are warranted in drawing conclusions from accepted premises on the basis of our

knowledge of logical laws. This is a fair assumption: we will discuss it in the next sections in more details, but its intuitive appeal should be clear. Carroll's regress challenges this assumption by purporting to show that this kind of warrant cannot ground the move from premises to conclusion. These two elements, taken together, lead naturally to a repudiation of the internalist assumption and to the espousal of a standpoint that regards our inferential practice as not being grounded on our possession of knowledge of logical laws.

Such a standpoint can take two forms. One can adopt a broadly externalist model of inference, according to which no internalist warrant is necessary to move from warranted premises to a conclusion. The externalist model construes warrant in terms other than explicit knowledge of logical laws. An externalist might cash out his notion of warrant by appealing to the idea that it is enough, in order to be warranted to draw an inference, to be a reliable agent, to arrive at the conclusion via appropriate methods, to be in the right circumstances, and so forth. However, someone who accepts the conclusion of the regress but wishes to resist the notion of an externalist warrant will embrace scepticism — namely, the view that there are no norms of inferential practice.

I will attempt, in what follows, to defend an internalist solution to the problem. The main strategy is to show that externalism and scepticism are not satisfying standpoints to adopt with regard to our inferential practice, and then to offer an internalist alternative.

In the next two sections, Quinean matters will be discussed in relation to this dialectic. Firstly, Quine's critique of Carnap's conventionalism about logical laws will be examined — we will see how Carnap's conception of logical laws is subjected to the regress. Then, Quine's own philosophy of logic will be assessed in the light of the considerations just reviewed. The situation in Quine's case, as we will see, is more ambivalent. However interpreted, though, it is unstable. The sections following

these Quinean considerations seek to resist the externalist and sceptical standpoints by exploring an internalist solution.

2.5.4   Quine's critique of Carnap

Quine mentions the regress in relation to Carnap in his 'Truth by Convention' (1936b). In the last pages of this paper, he discusses Carnap's idea that logical laws are conventions. As is well-known, Quine has no sympathy for such a view — one of the reasons being that conventions are, in Carnap's system, the cornerstone of the analytic / synthetic distinction. Directly connected to our present concern is this: Quine argues that the view according to which logical laws are conventions warranting inferences falls prey to Carroll's regress. Indeed, this should not be surprising. In Carnap's view, the conventions stating the truth of the logical laws are general and use schematic letters. (**W**) is a good approximation of what a convention stating the truth of all applications of modus ponens would look like in Carnap's system. According to Carnap, conventions such as (**W**) declare the truth of an infinity of statements — all the statements displaying the appropriate form. In Carnap's system, we need such a general convention since it is impossible to assert individually the correctness of every instance of modus ponens due to the infinity of statements displaying this form. If it were possible to assert individually the correctness of all instances of modus ponens, the difficulty would disappear. To return to our example, we would not need to proceed 'mediately' through the general convention (**W**) in order to draw the conclusion (C) and therefore there would be no regress — the further argument consisting of (a), (b), (c) would not be needed. The correctness of this particular argument would be declared true by convention in advance, maybe an item on a long list whose heading is 'Valid Arguments'. But as it is, in Carnap's system the conventions are what warrant any particular conclusion.

This detour via Carnap reveals Quine's diagnosis of Carroll's regress. According

to Quine, the regress starts when we appeal to the general logical law in order to warrant the move from the premises to the conclusion: "In a word, the difficulty is that if logic is to proceed *mediately* from conventions, logic is needed for inferring logic from the conventions" (1936b: 97). To put it in another way, the use of the convention (**W**) in Carnap's system is needed in order to infer, from the convention, the correctness of a particular application of modus ponens. If (**W**) is needed to infer (C) from (1) and (2), then it is also needed to infer the conclusion that the conclusion (C) is valid from the claim that: (a) (**W**) is correct and (b) the argument consisting of (1), (2) and (C) is an instance of modus ponens. It is not an option to hold that the law is required in one application and not in the other. It thus appears that Carnap's model is internalist. Quine's diagnosis, then, proceeds in two steps: if we understand (**W**) as being a convention (as it has to be, in Carnap's system, given the infinity of their instances) and if logic is to proceed mediately via those conventions, then regress ensues.

It is hard to dispute Quine's criticism of Carnap on this count. Indeed, Carnap's system, as understood here, does not have the resources to resist Carroll's regress. This is because the idea of a convention brings with it the idea of its being a warrant. What is a convention for if it is not capable of being used as a warrant? Isn't the *raison d'être* of a convention just that — being capable of acting as a warrant? Indeed, conventions are usually thought of as norms *grounding* our judgements, norms that we *follow*. If this feature is constitutive of conventions, then their capacity to act as warrants seems also constitutive. Moreover, since logic proceeds *mediately* from conventions, this model of logical knowledge is an internalist one. If this is true, then it shows that the very nature of conventionalism about logical laws contains the seeds of its vulnerability to Carroll's regress. Accordingly, an easy fix is out of the picture — I would go as far as to venture that *any* fix is bound to disappoint. Let us turn, then, to a different philosophy of logic.

2.5.5   Quine's philosophy of logic and the regress

Now what about Quine's own philosophy of logic? Does it afford the means to avoid Carroll's regress? Given that his diagnosis puts the blame, partly, on the fact that in Carnap's system logical laws are conventions stating the truth of an infinity of statements, it appears that it does not. Indeed, in Quine's logical theory, statements of logical laws are metatheorems and those need not be understood as conventions. However, a crucial part of Quine's diagnosis hinges on the need, in Carnap's system, to proceed *from* conventions *to* their applications. Viewed in this light, the metatheorem / theorem distinction fares no better. If the metatheorems are needed to warrant an application of a logical law (such an application being made to theorems expressing no generality), the regress threatens. In his logic books, Quine does write in such a way as to suggest that particular inferences do proceed via the metatheorems — as we have seen in section 2.4. The metatheorems establish the truth of an infinity of statements and they state the conditions under which statements are theorems. All the ingredients are apparently there for the regress. Despite the fact that the metatheorems are not understood as conventions, this set up is almost enough to generate the regress. The only additional element required is the inexorability of appealing to the metatheorems in order to warrant particular inferences. This additional element seems to be implicit in the very idea of the metatheorem / theorem distinction. If this is correct, Quine's philosophy of logic fares no better than Carnap's in relation to the regress of warrants.

On the other hand, consider this. Quine advocates 'naturalised epistemology', according to which epistemology is but a branch of psychology. Quine starts by making a distinction inside foundational studies between conceptual and doctrinal studies.[15] The conceptual side of a foundational enterprise (epistemology of logic in the present case) is concerned with clarifying and defining the central concepts of

---

[15]See Quine (1969a: 69-71) on this distinction.

the area of discourse in question. The doctrinal side is concerned with truth and justification, with warrant and certainty. A main tenet of naturalised epistemology is the abandonment of the doctrinal side — the epistemologist need not be concerned with notions such as justification, warrant, ultimate truth, certainty, and so on. All that is called for is an understanding of how we come to acquire belief — a chapter of empirical psychology. The Cartesian quest for certainty should be abandoned. Quine said in that paper that 'the Humean predicament is the human predicament'. Hume failed in his project of showing how certainty was possible, how we could avoid scepticism about our knowledge of a wide range of statements. In other words, he could not propose a satisfying doctrinal side to his epistemology. Quine's suggestion when talking about the human predicament is that such a doctrinal theory is bound to disappoint. This is why he put forward his naturalised epistemology:

> Epistemology, or something like it, simply falls into place as a chapter of psychology and hence of natural science. It studies a natural phenomenon, viz., a physical human subject. This human subject is accorded a certain experimentally controlled input — certain patterns of irradiation in assorted frequencies, for instance — and in the fullness of time the subject delivers as output a description of the three-dimensional external world and its history. The relation between the meager input and the torrential output is a relation that we are prompted to study for the same reasons that always prompted epistemology; namely, in order to see how evidence relates to theory, and in what ways one's theory of nature transcends any available evidence. (1969a: 82-83)

The point of quoting this passage in full is to see how talk of truth, justification, and warrant is simply not in the picture of this new conception of epistemology. Epistemology, for Quine, might be called a 'foundational' science but it does not provide any kind of foundation, strictly speaking, for knowledge.[16]

---

[16]For a fuller discussion of Quine's conception of epistemology, see Hookway (1988), especially

Let us now come back to the topic at hand. Given his conception of epistemology, Quine could be understood to be a sceptic about our norms of inferential practice. Indeed, what we are trying to do here is to provide a ground for our most fundamental inferential practices. But to search for a ground, for what warrants us to go from accepted premises to correct conclusions, is an attempt to do the impossible — it's an attempt to transcend our 'human predicament'. Looking for norms playing the role of warrants is a completely misguided project from the perspective of naturalised epistemology. We should merely look at what humans actually do, at the input they receive and the output they deliver. We recognise the truth of a set of premises (input); then we deliver, as output, a conclusion. All that can be said is that most individuals will be inclined to deliver the same output given similar inputs. And that's it — nothing more can be said, except, perhaps, pointing at regularities and trying to explain them naturalistically.

So there is an apparent tension in Quine's writings. On the one hand, he seems to be a traditional logician with a foundational, internalist view of logical laws with his metatheorem / theorem distinction. On the face of it, the metatheorems (logical laws) warrant the moves in the language-game of inference. On the other hand, Quine advocates, via his naturalised epistemology, scepticism about inferential norms. Where do matters stand? Is Quine inconsistent?

I don't believe so — at least not at this stage. There is a way out for the Quinean. She can hold that metatheorems — expressing the generality of logic — are a mere 'representation' or 'systematisation' of the way we actually reason, and as such we do not need to use them when inferring. Such an understanding of logical laws would block the regress, since there would be no need, in our example, to invoke a further argument having the modus ponens form in order to infer (C). The general logical laws would have an expository role, their usefulness being that they systematise how

Chapter 11. Stroud (1984) also contains a discussion of naturalised epistemology.

inferring is actually practised by us. According to Quine, this is consistent with naturalised epistemology. Indeed, immediately after the passage quoted above, he says that this new conception of epistemology allows for a "rational reconstruction, to whatever degree such reconstruction is practicable; for imaginative constructions can afford hints of actual psychological processes, in much the way that mechanical simulations can." So logic (understood as the systematisation of inferential practice, with rules, logical laws, and so forth), on this view, is a representation of our practice. It has no normative force whatsoever. It sole utility is to study psychological processes.

However, is this move available to a Quinean? Under a strict interpretation of Quine's writings, it is not. For sure, Quine, more often than not, writes as if logical laws are part of the web and *regulate* its logical interconnections. If general logical laws really have that regulatory role, it means that they are needed — in some sense — when performing a particular inference. It would be *because* of the metatheorems that such and such inferences are valid. In other words: if the logical laws have a regulatory role, that means, presumably, that they act as warrants.

Moreover, if we understand (**W**) in the example of last section as being a metatheorem regulating particular instances of it, and if we assign to the notion of 'regulation' a justificatory role, then we are espousing internalism and there is no way to escape Carroll's regress. It is enough for the regress to start that knowledge of general logical laws is required in order to draw inferences. Therefore, if we interpret some of Quine's passages to the letter and put his naturalised epistemology aside, his philosophy of logic cannot answer Carroll's regress.

It is, no doubt, possible to ignore Quine's own apparent understanding of the metatheorem / theorem distinction and offer a Quinean answer reconciling naturalised epistemology with logic along the lines previously mentioned. Under that interpretation, as we have seen, general logical laws — especially rules of inference such as modus ponens — are to be understood as theoretical representations of the

way we actually infer. I will call this the 'logic-as-representation' view. Logical laws, on this view, need not be considered as norms whose role is to warrant inferences. To infer in such and such a way is merely a practice ungoverned by any norms. A 'correct' judgement of logical consequence, on this view, is a judgement that will command assent by the majority of the speakers of a linguistic community. Since there are no general rules under which a given speaker has to proceed when inferring, any speaker is free to infer the way he sees fit. Strictly speaking, then, there cannot be an 'incorrect' statement of logical consequence due to the absence of norms. There are no logical laws in the web — no statements regulating the logical interconnections between statements. A 'devious' inferrer is someone whose deductions are at odds with what the majority of the speakers in his community would assent to.

How, then, should we understand the role of logic in such a picture? Logical laws would be representations of the way the community at large infers, but not rules that are followed. The immediate worry here is that this so-called theoretical representation does have a justificatory role. Someone pressed to explain why she infers in such and such a way could reply that the only norm she obeys is not a general logical law but rather a communal norm stated in terms of behavioural assent and dissent of particular instances of inferences. The community does, or would, agree on the correctness of this inference, and this is what makes it a correct inference. Logic, on the sceptical view, is a mere representation of the inferring practice of this community, and it has no normative role. But this doesn't seem to accord with practice. Indeed, when someone makes an incorrect inference, we do not merely tell that person: 'what you infer in this case is not correct because this is not what the rest of the linguistic community would infer given those premises'. No — we would normally tell him: 'look, what you have done is not right because you have not applied this general logical law correctly, and this is the correct way of inferring'. Upon realisation that he did not apply the logical law, two options are open to the

devious inferrer. Either he corrects his inference and acknowledges that, on the basis of the logical law, he should infer such and such; or he disputes the validity of the logical law. In both cases, the normative role of logic is conceded. It seems difficult to reconcile naturalised epistemology with a view acknowledging the importance of logical systems.

A stubborn Quinean who feels uncomfortable with the marriage of naturalised epistemology with the 'logic-as-representation' conception could bite the bullet and abandon the latter. That is, such a Quinean could suggest a radical scepticism, doing away with logical systems and logical laws altogether. Independently of the credentials of that sceptical stance, it is safe to say that it isn't a Quinean view anymore. Logic — understood as logical laws formulated in full generality — plays a prominent role for any self-respecting follower of Quine. More importantly, the view that there are no logical norms is an extreme one for non-Quineans as well as for Quineans. It would indeed remove the ghost of Carroll's regress (since there would be no laws such as modus ponens to rely on when making inferences), but at a considerable cost. If it were the correct view, then we should throw away all our logic books and eliminate logic courses from the philosophy curriculum. Inference being simply what we do, the function of logic books would become a mystery. It can't be replied, of course, that in logic books what we find are the norms that the community follows: this would not amount to scepticism.

To recapitulate. Quine claims both that logical laws have a regulatory role and that epistemology should be naturalised. But these two claims cannot be held together — naturalised epistemology implies a sceptical view about logical norms and the incorrectness of any 'doctrinal' studies (that is, studies having to do with justification and truth). This shows his position to be unstable. But any attempt to stabilise it leads to undesirable conclusions. Say a Quinean decides to remedy the situation by abandoning naturalised epistemology. In that case, the resulting philosophy of logic

is vulnerable to Carroll's regress, just like Carnap's. If, on the contrary, our Quinean decides to abandon the regulatory view of logic, the problem becomes: how to make sense of logic in that picture? How do we make sense of the logic-as-representation conception? It seems that a reasonable position, if one is to take that line, is to drop the idea of logic as mere representation. However, doing so leads to an extreme scepticism. Such a view is not Quine's, I believe. So on whatever way of looking at it, Quine's philosophy of logic does not have the resources to answer Carroll's challenge. However, let us now go beyond Quine and see whether sense can be made of the resolutely sceptical view.

2.5.6 Logical norms and scepticism

The question that Carroll's regress forces us to answer is this: given a non-empty set of premises and a conclusion, what is the nature of the warrant, if any, enabling us to draw the inference and say that the premises entail the conclusion? The regress shows conclusively that the warrant in question cannot be one formulated in terms of a further argument licensing the adoption of the conclusion in the light of the other premises since for this argument to properly *warrant* the conclusion, appeal to (W) must be made. The solution to Carroll's regress, accordingly, has to be that the addition of further premises licensing the move from the premises to the conclusion must be restricted in some way. There are three general forms such a solution might take. The first is to acknowledge the existence of logical norms but to deny that they must be appealed to when making inferences — this is externalism. The second is to advocate a sceptical stance and dispense with norms altogether. The third is to allow a normative role for a restricted class of logical laws but to deny that they have to figure as further premises forming an additional argument; it corresponds to a modified internalism. This section investigates the prospect of scepticism whilst the next section will be occupied with externalism. The remainder explores the possibility

of a modified internalism.

Dispensing with norms is to embrace the sceptical stance, namely that when we infer, there is no warrant in play — we do not, and should not, appeal to logical laws when inferring. To espouse scepticism is to repudiate the apparent fact that there are *norms* that we follow in our inferential practice. This strikes me as a highly revisionist claim. On this view, logic (by that I mean logic as exposed in logic books, logical systems, and so forth) is a mere representation, a 'rational reconstruction' of inputs and outputs. This reconstruction has absolutely no normative force: we can't use the reconstruction as a warrant, otherwise we wouldn't be embracing scepticism anymore. Indeed, scepticism does seem to be, on the face of it, committed to repudiating norms. What is the place of logic in such a picture? Is there a place for it? Here we seem to face a dilemma. The first horn has it that logic is a mere representation, a rational reconstruction. But it has to claim that this representation has no normative force. This strikes me as implausible. It's comparable to claiming that maps represent a country (say) while holding that the map has no normative status. Of course not: the whole point is to use the map as a normative (if conceivably flawed) tool. Similarly, the point of devising a representation of how we infer is to guide our inferential practice. This is what logic books are for! The point I'm making can accommodate the idea that logic represents the way we (rational humans) actually infer. The norm, then, is the way we actually infer. If that much is conceded, the sceptical standpoint is abandoned: when inferring, we appeal to norms such as (**W**).

Before turning to the other horn of the dilemma, I want to examine a possible reply by the sceptic. Here she could claim that we use the representation as a guide, but we don't have to. It is not necessary for correct inferential practice to appeal to norms. When we do refer to logic books, we are merely taking a shortcut instead of going through the process ourselves. What makes an inference correct is not its compliance with norms, but something not normative (whatever that may be). To

come back to the map analogy, it is not the map that makes true which road to take, but rather the actual road system, out there in the world. But it is precisely at this point that the map analogy threatens to break down: what corresponds to the 'right' road, out there in the world, when we talk about inference? This reply by the sceptic, as it turns out, amounts to an expression of externalism about inference, and we will come back to that in a minute.

The other horn of the dilemma is to repudiate the logic-as-representation view and hold that the function of logic books and logical systems is by no means clear. The problem here is immediate: what have we been doing all these centuries? Were we just *wrong* in devising logical systems? Were we wasting our time? Obviously not. What are students learning in logic classes, if not norms of correct reasoning? It seems that if we are to be consistent sceptics about the normativity of logic, we should follow a paraphrase of Hume's advice and 'commit logic books to the flames, for they contain nothing but sophistry and illusion'. If this is not radical revisionism, I do not know what is.

The dilemma, then, is this. If the sceptic endorses the logic-as-representation view, then it's hard to avoid the recognition of the normativity of logic. If, on the other hand, the sceptic decides to do without logic altogether, the view is extreme. In a nutshell, the sceptical conclusion just seems to be inconsistent with the way we infer, since we do, as a matter of fact, reason according to rules of inference and we justify our inferences on the basis of them, etc. It is committed to the view that such talk is illusory, and this is, I believe, intolerable. I acknowledge that the sceptic, however, may just be unmoved by such arguments — the considerations just advanced give us reasons to *resist* the sceptical conclusion, but they do not *refute* it.

2.5.7   The externalist proposal

Let us now turn to externalism. The idea would be to claim that we infer the way we do because this is what the world requires of us. Faced with appropriate premises whose truth we recognise forces the acceptance of the conclusion without any process of self-reflection involved in the way the regress demands — without, that is, any 'internalised' warrant being necessary. Indeed, the regress goes through only if one has to recognise that the premises fulfil a certain condition, and then, on one's recognition that they do fulfil those conditions, one is licensed to infer the conclusion. But if it is claimed that inferring the way we do is simply a matter of recognising the truth of the premises and that the conclusion will 'force' itself on us because of externalist considerations, such an appeal to the required condition (the warrant) doesn't play any role — it becomes superfluous. Thus, upon's one recognition that all men are mortal and that Socrates is a man, the conclusion that Socrates is mortal forces itself on us — in terms of input and output, given the input of the premises, the natural output is the conclusion. Of course, the problem here is to make sense of the metaphor in play when talking about the conclusion 'forcing' itself on us. To make matters clearer, think of the parallel here with a famous case, that of the recognition of chicken sexes. Just as it can be said that recognisers of chicken sexes are able to do so without being able to justify their capacity by appealing to a criterion, one could hold that we infer correctly without appealing to a rule of inference — all that matters is that the rule *is* valid. Exposure to chickens over a period of time will somehow instil the capacity to distinguish between their sexes: similarly, exposure (so to speak) to appropriate premises will, for any competent speaker, prompt the capacity to infer correctly — and that's it, nothing more is required, strictly speaking.

The immediate problem with such a solution is that there is an apparent dissimilarity between the two cases. It is uncontroversial that there is a fact of the matter about chicken sexes — the controversial externalist thought is the offered explanation

of the capacity to discriminate between the sexes (no need for internalised warrant), not whether there is a fact of the matter or not about the sex of chickens. But if we push the parallel further, it appears that in order to be able to offer an externalist explanation of our capacity to infer correctly, there has to be a fact of the matter about logical inference. The externalist offers an explanation of the route between an external fact and the capacity of a cognitive agent to recognise that fact. In the case of inferential practice, a proper externalist explanation will have to proceed along these lines — namely, that the inferring agent is in contact with something over and above the recognition of the truth of the premises, and that this something is what enables the agent to draw the inference. This 'something' would have to be a valid rule of inference. But this seems to involve an ontological commitment which is puzzling. What is it that we recognise when we infer correctly? What are we responding to, and how? Are we somehow in contact with the rules of inference — the same way that the chicken sexes discriminator is in contact with the relevant features? The externalist thought about chicken sexes has some credibility since the ontological issue is not in dispute. No one disputes the existence of chicken sexes. But what is external about a rule of inference? What does it mean to say that we are in contact with rules of inference and that we 'unconsciously', as it were, track these rules? These are wide-ranging questions. Even if it were possible to answer them, they would lead us too far from the topic discussed. It suffices, at this stage, to identify these questions. A more plausible solution to Carroll's regress will be offered in the next section, a solution not embodying such a strong ontological commitment.

In the last two sections, I have reviewed a general claim that has to be endorsed by the proponent of scepticism (the denial of norms) and an externalist account of our inferential practice. It is notoriously difficult to counter-argue in a decisive manner against the externalist and the sceptic. The situation here concerning scepticism is the following: any form of the scepticism appears to be committed to denying

that there are norms that we follow in our inferring practice. Scepticism just seems to be inconsistent with the way we infer, since we do, as a matter of fact, reason according to rules of inference and we justify our inferences on the basis of them, etc. It is committed to the view that such talk is illusory, and this is, I believe, intolerable. If one prefers to embrace externalism, the difficulty is to make sense of the non-internalist warrant that we follow when inferring correctly. No matter how difficult — if possible at all — it is to actually *show* that all externalist and sceptical standpoints do not work, there still is a way to argue for an internalist solution. The idea is to keep in mind how the sceptical thought does not accord with what seems to be, prima facie, how we, as a matter of fact, infer, and then to suggest a non-externalist account which is not prey to Carroll's regress. If it is possible to show: (1) the plausibility of an account of our inferring practice doing justice to the idea that knowledge of (**W**) and the like is required to draw inferences, and (2) that this proposed account does not fall into the regress, then, I believe, the non-internalist standpoints will prove to be much less attractive.

2.5.8 Constitutive norms

There is another line of thought available as a response to Carroll's regress which is more congenial to an internalist outlook. It draws on Davidsonian ideas about the nature of practical reasoning. In order to see how those Davidsonian ideas can be exploited in the present context, it is necessary to distinguish between rules of inference that are *constitutive* of our practice of inferring and those that are not constitutive. The focus has so far been on modus ponens and not on other logical laws. This very fact points towards a special feature of this logical law. I will argue in what follows that if a satisfying internalist proposal is to be possible, then the constitutive role of modus ponens has to be recognised.

Let us assume, for the sake of the argument, that logical laws have a normative

status in the sense that once we accept (implicitly or explicitly) those laws, we are bound to infer according to what they prescribe — pending their possible revision. It is not incoherent — even if it might possibly be a mistake — to abandon the law of excluded middle or (more radically) the law of non-contradiction. But as long as we accept those laws, we are bound to infer according to them. In this sense, logical laws are norms. Modus ponens is also a norm of inferential practice, but it is also a constitutive norm in the sense that the very practice of inferring presupposes modus ponens. Carroll's regress, in my opinion, reveals this interesting fact about inference.

To see why, let us take this simple inference, an inference *not* using modus ponens:

(20) Socrates is mortal and Aristotle is Greek.

(C*) Aristotle is Greek.

On an internalist account of our inferential practice allowing for justificatory norms understood in the usual sense, this inference is subject to the regress. Indeed, to be able to draw the conclusion (C*), you must be credited with the knowledge that the argument above displays a valid argument-form, that is, you must know that the rule for &–elimination is valid. How does this knowledge warrant you to draw the conclusion (C*)? Well, this warrant must be represented like this:

(a*) (**W***): If an argument displays the logical form of the &–elimination rule, then it is valid.

(b*) The argument above ((20), (C*)) is an instance of the &–elimination rule.

(c*) Therefore, the conclusion (C*) — Aristotle is Greek — can be drawn.

But this is another inference, an inference justifiable only if one is warranted to accept modus ponens. So in order to be warranted in drawing the conclusion (c*),

you must be credited with the knowledge that:

(i*) (**W**): If an argument displays the logical form of modus ponens, then it is valid.

(ii*) The argument above ((a*), (b*), (c*)) is an instance of modus ponens.

(iii*) Therefore, the conclusion (c*) — that the conclusion (C*) can be drawn — can be drawn.

And so on. Whatever rule of inference is used in the original argument, an ultimate appeal to (**W**) must be made. This is because no matter which rule is appealed to in the first place, it is possible to construe an argument having the form (a), (b), (c) or its starred equivalent. Any such argument will be warranted, ultimately, on the basis of modus ponens since for any instance of any logical law, to know that the conclusion follows from the premises requires grasping an argument which has a modus ponens form — namely, it requires two thing: (1) that *if* the argument displays such and such a logical form, *then* it is valid, and (2) that the argument displays such and such a logical form.

In the end, once appeal to (**W**) is required, the regress is unavoidable. It is because all rules of inference presuppose, for their warrant, a further argument having the modus ponens form that Carroll's regress will go through for them. More precisely, it is because any explicit appeal to any rule of inference when justifying a particular move from premises to conclusion is itself dependent on another rule of inference — modus ponens — that there is a regress.

One could say, then, that modus ponens is constitutive of our practice of inferring in the sense that inferring requires a warrant involving this logical law. In other words, no warranted inference is possible for us without using modus ponens. In contrast, the law of excluded middle does not seem to be constitutive, in the sense just mentioned, of our practice of inferring. Indeed, it is not required when inferring

using the law of conjunction or the law of modus ponens, as can be seen from the examples of the previous three sections. It has to be acknowledged that nothing in this argument excludes that there might be other constitutive logical laws. It is enough for our purposes that there is one such law.

2.5.9 An internalist proposal

Now how does the distinction between constitutive and non-constitutive logical laws help us in solving Carroll's regress? Taking my lead from Davidson's ideas about akrasia, I will argue that we can draw a parallel between his proposed solution of this problem of practical reason and the present difficulty. Akrasia occurs when an agent's action does not conform to what she recognises as the right action to take. For instance, someone is offered a cigarette. She arrives at the conclusion that since smoking is bad for health, causes cancer, etc., the right action to do is to refuse the offer. Nevertheless, she accepts the offer and lights the cigarette. She acts contrary to what her own practical reasoning prescribes her to do — she is an akratic agent. How can we explain the phenomenon of akrasia? It seems irrational to accept the offered cigarette, the agent recognises that it is irrational, but still she accepts it. The situation could be schematically put like this. The agent's reasoning could be pictured thus:

(PR1) Cigarette causes cancer.

(PR2) I am being offered a cigarette.

(CPR) Therefore, I should refuse the offer.

However, she does not do what her practical reasoning tells her to do: she acts contrary to it — and she is fully aware that she does. How should we explain that? Akrasia is a challenge for anyone interested in clarifying the nature of practical rea-

soning, just as Carroll's regress is a challenge for those interested to get a clear view of the nature of our inferential practice.

According to Davidson (1970; 1982), the explanation is the following. The agent is simply acting irrationally. The irrationality of the act resides in the fact that the agent ignored a central maxim of practical reasoning, that of the principle of continence:

Principle of continence: One ought to act on what one regards as the best option, everything considered.

This principle, in the case we are interested in, prescribes acting in accordance with (CPR) — this is what she regards as the best option, everything considered. But she does not act on it. She is acting irrationally. Davidson's analysis contains many twists and subtleties that need not concern us here, but let us focus on the status of the principle. In practical reasoning, one cannot choose to accept or to reject the principle of continence since any instance of practical reasoning presupposes use of the principle. The very idea of 'consciously' rejecting the principle would involve use of the principle itself. Therefore, one cannot rationally reject the principle — it cannot be *rational* to abandon the principle since it is impossible to give a reason for the rejection. This is the crux of Davidson's analysis of akrasia: the akratic agent is *irrational* since she acts contrary to the principle of continence. Moreover, and crucially, according to Davidson the principle of continence "is not just another premise to be accepted or rejected. It is constitutive of the concept of a creature with propositional attitudes capable of practical reasoning" (Davidson, 1999, p.404). So the principle is *constitutive* of practical reasoning and is no ordinary premise to be added freely when justifying a particular judgement. Arguably, one of the reasons why it cannot be added as a further premise has to do with considerations similar to Carroll's regress: adding the principle to justify an instance of practical reasoning

would, it seems, generates a regress similar to Carroll's.[17]

Similarly, we can say that since modus ponens is constitutive of our practice of inferring, it is impossible to reject it on rational grounds. The fact that it is impossible to rationally reject it reminds us of its special status: in contrast, it is possible to reject the law of excluded middle without being irrational. The validity of excluded middle can be discussed on rational grounds in a way that is not possible for modus ponens. Indeed, one can reason without using excluded middle as a logical law — intuitionists do it all the time. However, no one sharing our inferring practice — including intuitionists and others — can infer without using modus ponens. To dispute modus ponens is not rational since any discussion about its validity presupposes its use. Since modus ponens is constitutive of inferential practice, it cannot appear as an ordinary warrant when justifying a particular inference. A warrant, it could be argued, is something that can be debated. But the validity of modus ponens cannot be debated, as we have just seen. So when justifying an inference, modus ponens is not merely another premise to be added in support of the inference. It is still irrational not to reason according to modus ponens — a virtue of my proposal, as will we see, is precisely that it manages to explain the rationality of reasoning according to modus ponens without falling into the regress.

If this is right, then Carroll's regress is blocked: (W) is not to be added as a further premise fulfilling a justifying role. Remember that in the exposition of the regress, it started when knowledge of (W) had to be added as a further premise, namely the premise (a). But if modus ponens, because of its constitutive status, does not figure as an extra premise, then there is no regress. What happens when we are faced with an argument displaying the logical form of modus ponens is this: we recognise that the argument is of a certain form, and we draw the conclusion — and that's it, nothing more needs to be said.

---

[17]See Blackburn (1995) and Lazar (1999) for a discussion of this point.

However, there is a worry to be addressed: doesn't the proposed solution boil down to externalism, or worse, to scepticism? Indeed, if the recognition of the truth of the premises is sufficient for the drawing of the conclusion, then it appears that knowledge of (**W**) is not necessary — and this seems to be exactly what a non-internalist would want to say. I wish here to give two reasons to defend this criticism of the solution suggested here. First, modus ponens as understood in the proposal does have normative force — so the solution is resolutely non-sceptical. Indeed, one who does not reason in accordance with modus ponens is deeply mistaken. One *has* to infer according to this law, that was the whole point of showing its constitutive status. It is a norm that has to be followed, but which cannot be given, explicitly, as a warrant such that it would appear as a regular premise. If someone asks me why, faced with an argument with the modus ponens form, I draw the conclusion, I can point to modus ponens — but if my interlocutor does not understand and keeps asking for justification, at some point I can only say 'but can't you see?'. Most of the time (probably always, in fact), pointing to the modus ponens schema *is* sufficient for the interlocutor, he is likely to react as if saying 'oh yes, right, I get it'. The fact that the interlocutor is being shown the law and nods his understanding shows that modus ponens is a norm. But to appeal to it in the regular manner will lead to a regress if the interlocutor is stubborn — at some point, the step has to be made without adding (**W**) as a further premise. But that does not mean that modus ponens has no normative force whatsoever. So my proposal is to be distinguished from scepticism.

Secondly, externalism concerning logical laws is committed to the claim that knowledge of logical laws is never necessary. The proposal I am putting forward does not entail that. Indeed, that proposal has it that logical laws other than modus ponens keep their ordinary, fully internalised status. A virtue of this proposed solution is that it 'saves' the ordinary status of other logical laws. Very often we debate about logical laws: we put them as premises in arguments, we discuss them, try to

justify them, etc. This solution arguably excludes this kind of debate where modus ponens is concerned, but it leaves that possibility open for the other laws. They act as ordinary warrants, nothing is changed. The only difference is with the constitutive logical law. In its case, it is true that something seems conceded to the externalist — namely, knowledge of the norm followed need not be spelt out in the way the regress demands. But does that mean that knowledge of the norm is not required? Perhaps. But let's assume, for the sake of the argument, that under my proposal something *is* conceded to the externalist, namely: the move from premises to conclusion in the case of arguments displaying the modus ponens form is to be explained by a purely externalist account. If this is the case, then it is still a far cry from the externalist excesses criticised previously, where knowledge of all logical laws becomes superfluous.

2.6   Shapiro's discussion of Quine

Stewart Shapiro, in his paper 'The Status of Logic' (2002), also discusses the problem of locating the principles of inference in the Quinean web of belief. Shapiro's main concern is with the a priori: he addresses the question of whether logic is a priori or not. Since it seems to be the case that if logic is a priori then it is also necessary, Shapiro's claim — that even the Quinean has to accept that (part of) logic is a priori — is relevant to our topic.[18] Before examining Shapiro's argument, however, we need to introduce one of the notions he develops in the paper.

Shapiro points out that, as remarked in previous sections of the present Chapter, the principles of inference governing the logical connections in the web cannot be statements in the web itself. Taking this problem seriously, Shapiro suggests a way out for the Quinean by pointing to an equivalence between validity and logical truth. The key idea is that an inference from a sentence p to a sentence q is *valid* if, and

---

[18]Note that this connection between a priori and necessity does not always hold. For instance, it is a priori that 'I am here' is true, but it is not necessary that I am here.

only if, the sentence 'if p then q' is a logical truth. That is, let's say that q follows from p in the web. It seems that, as we have seen, there is no conceptual space in the web for such statements of logical consequence (statements connecting two (sets of) sentences with idioms such as 'logically follows from', 'is a logical consequence of', 'the inference between A and B is valid'). Quine cannot give an account of this kind of statement, mainly because of the regress discovered by Carroll. Shapiro's idea offers a way out by reducing the problematic statement of validity (impossible to locate in the web) to a logical truth which is part of the web. The reduction is done through what he calls the transfer principle, which has this form:

(**TP**) Inference $(\Gamma \vdash \Theta)$ is valid $\equiv$ the sentence '$\Gamma \to \Theta$' is logically true.

The transfer principle gives us a means to reduce talk about validity (logical consequence) to talk about logical truth. The notion of logical truth, in turn, is explained away by Quine's own substitutional account. Quine claims that his account of logical truth and of logical consequence does not involve modal notions, thus completing the reduction and, in effect, eliminating any modal notion. Quine's account of logical truth and logical consequence — effectively another manifestation of Error-theoretic Irrealism — is examined in Chapter 4.

The transfer principle gives us necessary and sufficient conditions for an inference to be valid. Let's illustrate the workings of the transfer principle by way of an example. Take the familiar valid inference:

(1) If Socrates is a man, then Socrates is mortal

(2) Socrates is a man

(C) Therefore, Socrates is mortal

So what we have is a left-hand side looking like this: 'the inference from (1) and (2) to (C) is valid'. This statement is in turn true by virtue of a logical law of inference — modus ponens — which is in effect a metatheorem. This logical law is the problematic statement, the one seemingly impossible to locate in the web. This statement is in effect a statement of logical consequence, one — to anticipate on what follows — of the form W studied by Wright in the next Chapter. Statements of this form regulate the logical connections in the web, i.e., they describe how to move from sentences to sentences, and this is why they are problematic. Shapiro's move is to reduce such statements to something like this (keeping with our example): 'the sentence "if (1) and (2), then (C)" is a logical truth'. Now this statement does not say anything about logical consequence; it is couched in terms of logical truth. It is about a particular statement in the web, not about how some statements relate to other statements. What the transfer principles do, in the final analysis, is to reduce the consequence relation (the relation of logical implication, holding between statements) '⊢' to the material implication (holding between the content of statements) '→' and add the predicate 'is logically true'. [19]

Let us note that the reducing sentence is part of the web since logical truths are neither metatheorems (difficult to locate in the web) nor statements of logical consequence (subjected to Carroll's regress). It should go without saying that Quine would applaud such a move. Reduction of a problematic notion to an unproblematic one is a manoeuvre which is constantly done in philosophy, and Quine himself certainly can be counted as a reductionist in many respects (*Word and Object* can virtually be read as a book performing mainly reduction after reduction — or even more radically, elimination after elimination). However, it should not be forgotten that even the reducing sentence is true in virtue of a metatheorem, since it is a logical truth.[20]

---

[19]Quine is aware of this equivalence between the relation of implication and logical truth. See his (1940: 5, especially p. 28).

[20]I will argue in Chapter 4 that Quine's account of logical truth and logical consequence is irre-

As Shapiro himself points out, the transfer principle limits the range of what is to count as an inference and what is to count as a logic. For instance, it is impossible to reduce infinitary inferences — such as the ω-rule for arithmetic — to a Quinean logical truth since such inferences are not reducible to sentences.[21] But such inferences are clearly valid, at least from an intuitive point of view. How can a Quinean make sense of that? Another case in which the transfer principle fails is when the logic under the consideration does not have a deduction theorem (such a theorem has the form 'if $\Gamma \vdash \Theta$, then $\Gamma \to \Theta$', that is, the left-to-right direction of the transfer principle). In some modal logics, Shapiro points out, the deduction theorem does not hold. Of course, Quine dislikes modal logic: but he certainly does not require, as an a priori criterion for a correct logic, a deduction theorem. In any case, Quine is a staunch advocate of classical logic.

But let us ignore these problems for the time being, since the prospect of such a reduction would represent an important advance on the issue of locating validity in the web. Shapiro then considers the status of the transfer principles (that is, statements instantiating the form (**TP**)). Are they problematic for a Quinean? Put in the framework of the next Chapter, the question is: can they be treated as mere hypotheses without fear of regress? To treat the transfer principles as hypotheses means to be able to envisage the possibility of revising them. If the Quinean can make a convincing case to that effect, she seems to have found a definitive way out of the problem we are investigating in this Chapter, *i.e.*, to make sense of validity (logical consequence) while espousing Error-theoretic Irrealism. If she cannot show that the principles are revisable, then, similarly to Wright's point in the next Chapter,

---

ducibly modal. If my arguments in that Chapter are correct, it follows that the attempt to reduce what is supposedly the irreducibly modal (according to Wright, as we will see in the next Chapter) and impossible to locate (according to the argument in the previous sections) statements of logical consequence to a less problematic statement of logical truth via the transfer principles is bound to fail before it starts. But let us here assume that Quine's treatment of logical truth is, as he thinks, purged of modalities.

[21] We shall have more to say about the ω-rule in Chapter 4.

it will show that the appeal to the transfer principles cannot save the Quinean from the charge that his picture of knowledge makes an implicit appeal to modality.

To say that the transfer principles are revisable is tantamount to say that after all, logical truth and validity do not coincide. That is, to revise a transfer principle is in effect to recognise that something is a valid logical inference but cannot be reduced to a logical truth, or vice-versa. So we are left with a 'valid' logical inference whose status is not explained by the reducing account outlined above, and we are back to the problem of making sense of validity. Or conversely, we have a logical truth stating an implication which is not capable of being transformed as a statement of validity: but then, in which sense is it a logical truth? It looks as if the Quinean cannot hold that the transfer principles are revisable. They would therefore be unrevisable and a priori: that is, necessary. This should come as no surprise. Remember that the equation of validity with logical truth is a provable theorem in some logical systems. This assimilation of validity to logical truth only works provided that it is possible to prove that it holds. The correctness of the transfer principles, therefore, depends on the validity of a certain proof. The proof is a metalogical one, or metatheoretic, and now the question becomes: what is the status of that proof? The problem is similar to the one encountered in previous sections, where doubts were raised about the intelligibility of a logical metatheory in the web.

Interestingly, Shapiro defends Quine and holds that the unrevisability of the transfer principles might not be as problematic as it appears in Quine's philosophy. Shapiro maintains that the transfer principles are fundamentally a priori. A sentence is fundamentally a priori, for him, "if it has an a priori warrant, and the reliability (or legitimacy) of this warrant is a priori, and the reliability of that is a priori, etc. — with the process continuing until we come on something that needs no warrant" (Shapiro, 2002: 11). He goes on to argue that Quine needn't be worried about the fact that his own system requires merely a *weak* set of fundamentally a priori logical

principles since the conception of the a priori that he was attacking in 'Two Dogmas' was much more embracing. So the idea seems to be that what is crucial here is a matter of degree: whereas the logical positivists needed a lot of a priori principles, Quine needs only a few of them in order to avoid regress. This restricted set of principles not being enough to cover the needs of the logical positivists, Quine can embrace them as a priori without selling out his argument in 'Two Dogmas'. However Shapiro (2002: 11) also notes that "(t)he notion of fundamental a priority is of a piece with traditional foundationalism". So Quine, forced to embrace a (weak) set of fundamentally a priori principles, is best viewed as a traditional foundationalist after all. If Shapiro is right, Quine's main target is not foundationalism as such, but merely its logical positivist version.

But would that be congenial to Quine? That is, would Quine be happy to be viewed as a traditional foundationalist? Many writers interpret him as attacking precisely traditional foundationalism. On this reading, Quine's attack on the logical positivists serves as a typical example of the kind of views he is ultimately challenging, namely traditional foundationalism. This latter reading gives wider scope to Quine's attack, and explains why it has been so influential. If Quine were, on the contrary, to be interpreted as attacking only the logical positivists, it would be hard to make sense of the tremendous influence of his argument. Quine challenges — as seen in Chapter 1 — the a priori / a posteriori distinction in 'Two Dogmas' in a way that leads one to believe that he is criticising the viability of the distinction as such, and not only as appropriated by the members of a philosophical movement. It is uncontroversial, however, that his attack applies and undermines logical positivism. The point I am trying to make is simply this: Quine is doing much more — he is in effect attacking the notion of a priori knowledge by espousing holism. Aware that the syndrome of the a priori is not completely dispelled by his critique of the two dogmas, he goes ahead and shows how we can do without it via holism. It is indubitable that he attacked

only a priori knowledge as manifested in the logical positivist's writings, but we must not lose sight of the fact that an alternative picture of knowledge in which the a priori plays a central role would certainly not be to Quine's taste, even if he did not put forward arguments against it. Even if the realm of the a priori which he would be forced to embrace would be restricted compared to the logical positivists, he would still owe us an account of a priori knowledge — but that would be inconsistent with his concept of naturalised epistemology. In a nutshell: the fact that Quine argues for the untenability of the logical positivist's version of a priori knowledge does not entail that he would accept another conception of the a priori.

Another problem with Shapiro's interpretation has to do with the conception supposedly advocated by Quine. There is first the question of how many principles will be a priori. Are we talking about a complete logic here, with axioms and rules of inference? Or should we rather have in mind only a few principles? The first option raises the worry that Quine's conception will be indistinguishable from the logical positivist's conception of the a priori. The second option brings to mind the worry that whilst the set of a priori principles will be smaller than the positivist's, this new set will be subjected to Quine's own criticism of the a priori. Another worry with this option is: how are we to determine these principles without developing a full logic? Can we have only a few logical principles without a full-blown logic with axioms and rules of inference? This is crucial, since if talk of 'a few principles' commits one to talk of a full-blown logic, then the advantages of restricting the realm of the a priori seem to fall apart.

To restrict the realm of the a priori also raises the question of how to determine which principles are necessary and sufficient to regulate the logical interconnections in the web. One suspects that modus ponens and a few rules regulating negation will be part of the set. But what about the rules governing conjunction? Should we include them? And if this minimal set of rules entails a lot of other rules, in what

sense can it be said to result in a sort of 'minimal' logic? The spectre of inflating the minimal set of principles to a full-blown logic appears again.

2.7 Conclusion

In this Chapter, we have studied Quine's holism and its relation to error-theoretic irrealism. After disentangling various notions playing a key role in Quine's picture of knowledge and its relation to his attack on analyticity, the question of how to locate logical laws in a holistic theory of knowledge was addressed. By studying the question via Carroll's regress, we have seen that Quine's holism does not have the resources to contain a satisfying philosophy of logic. Moreover, it was pointed out that Carroll's regress does not merely create a puzzle for holism. Indeed, the regress raises deep questions about the prima facie justificatory role of logical laws. The problem affects any perspective in the philosophy of logic acknowledging such a role to at least some logical laws. Since this justificatory role seems to be a normal feature of our understanding of inference, it appears that an argument purporting to show that logical laws do not and cannot have a normative status prompts an urgent need for a solution. After studying various possible solutions, I took a lead from Davidson's ideas on akrasia and suggested a possible way out. The main advantage of the proposed solution is that it retains the justificatory role of logical laws without falling prey to the regress. Finally, we studied Shapiro's attempt to reconcile the spirit of Quine's holism with an acceptance of Carroll's regress. His solution commits holism to making a restricted place for a priori knowledge. We concluded that such an attempt would still leave unexplained the status of this kind of knowledge; in particular it was pointed out that it is not clear whether holism has the resources to accommodate it. The next Chapter will be devoted to examining a well-known argument attacking Quine's philosophy of logic from a different perspective. This will, once more, lead us to look at issues going beyond the scope of Quine's scepticism

about logical necessity — issues forcing us to think hard about the very foundation of logic and inference.

CHAPTER 3

WRIGHT'S ARGUMENT

3.1   Introduction

In the last Chapter, I argued that Quine's web of belief picture couldn't give a satisfying account of the logical interconnections inside the web. The problem raised is an internal one: it has to do with the logical structure of the web. Quine's holism does not have the resources to explain the crucial role of the laws governing the relations between the sentences in the web and to explain the relation between the metatheory and the rest of the web. Wright, in his well-known paper on logical necessity (1986), also argues that the Quinean picture suffers from a problem having to do with the role of statements of logical necessity. His argument is intended to show that contrary to what Quine holds, the web of belief picture makes an implicit use of statements of logical necessity. That is, Quine claims he can do without logical necessity; Wright argues that the very intelligibility of Quine's picture requires logical necessity.

This Chapter examines Wright's argument. Section 3.2 explains the argument in detail, and in the following section a Quinean objection to Wright's argument is studied and rebutted. I then appraise what exactly is establish by Wright's argument. As we will see, there are doubts as to whether the argument, even if valid, establishes what it purports to establish, namely the necessity of logical necessity. Sections 3.4 and 3.5 answer these doubts by showing how the gap in Wright's argument can be filled.

3.2   Wright's argument

Wright tells us what happens in Quine's picture when a theory is faced with a recalcitrant experience. The Quinean picture, according to Wright, goes as follows. There is a theory, call it Θ, from which we can derive conditionals of the form $I \rightarrow P$, where I stands for a set of initial conditions and P for a prediction. Let us call such statements *observational conditionals*. $I \rightarrow P$ is obtained from Θ via a logic L. We

thus have, corresponding to each derived statement $I \rightarrow P$, another statement, W, which describes the derivation in Θ of the observational conditional. Such statements have the form '$\Theta \vdash_L (I \rightarrow P)$'. It says, in effect, that given a basis (a theory) and a logic L along with rules of inference, such and such theorems follow (in this case, observational conditionals). W has exactly the form $(\Gamma \vdash_L \Theta)$, a statement we would be inclined to say is logically necessary. (See Introduction for more on $(\Gamma \vdash_L \Theta)$).

According to Quine, the main virtue of a theory is that of maximising the fit between the theory and the data — in this case, since Quine is an empiricist, the data consist, presumably, of sentences that we are inclined to assent to on the basis of experience.[1] In the holistic picture of knowledge, the scientist is constantly revising her web of belief in the light of experience (more precisely, in the light of statements we are inclined to assent to on the basis of experience). She tests her observational conditionals on the tribunal of experience and modifies the web accordingly. The idea is clear: when the theory does not sufficiently fit the tribunal of experience, we should change — modify — the theory so that it maximises fit with the data.

Wright explains in detail what exactly is going on when one is making changes in the web. The main impetus for making such changes clearly resides in the occurrence of recalcitrant experience. Let us take an arbitrary observational conditional, obtained from an arbitrary theory. A recalcitrant experience E is an experience inclining us to dissent from P but to assent to I. That is to say, the initial conditions described by I are obtained but the prediction P does not occur — something else happens. What exactly is involved in describing an experience as recalcitrant? Let us construct an example of a recalcitrant experience, by way of a simple theory.[2] The theory Θ consists of, amongst other things, the following 'eternal sentence': 'Copper oxide is

---

[1]Note here that we can't say that the data consists of perceptual experience per se. In the web, there are only sentences. Saying that the theory confronts the 'pure', unconceptualised perceptual experiences would be to fall into the myth of the given. See Sellars (1956). Thanks to Stewart Shapiro for drawing my attention to this important point.

[2]I borrow the example from Quine himself. See *Word and Object*, pp. 10-13.

green'. An observational conditional, derived from Θ, goes as follows: 'if two test tubes containing, respectively, substance A and substance B are mixed, then we get a green stuff called copper oxide'. The phenomenon of recalcitrance is to be imagined as follows. The scientist prepares the tubes, mixes them, and has an experience E to the effect that the stuff does *not* turn green. In other words, to stick to Quinean terminology, the scientist dissents from the sentence 'the stuff is green' — her stimulus response to this sentence is one of dissent. But the theory predicts that the stuff would be green. So there is an experience which is recalcitrant.

It is clear that the recalcitrance of an experience is a function of many factors, including the theory Θ. E is recalcitrant only because we are working under the framework of a given theory, such a theory implying the observational conditional. An experience is not recalcitrant per se. It is recalcitrant because it contradicts an observational conditional. The scientist works under the hypothesis that her theory is true, and that the observational conditional is a correct prediction. The sentence assented to on the basis of experience E is also held to be veridical. But one cannot hold all of this to be true at the same time, since the sentence assented to on the basis of E (that is, $(I \& \sim P)$) contradicts the observational conditional $I \to P$. Unless one is prepared to accept such basic and blatant contradictions, one has to change something in the theory. To be more precise, recalcitrance requires three main elements. First, there has to be an observational conditional $I \to P$, obtained from Θ via a logic L. Without such a conditional, there would be nothing to be recalcitrant about. Second, there has to be an experience E. And thirdly, this experience must prompt assent to I and $\sim P$. To recap, we need for recalcitrance:

1. $W\colon \Theta \vdash_L (I \to P)$.

2. An experience E.

3. E prompts assent to $(I \& \sim P)$.

Strictly speaking, it is the conjunction of (1) – (3) which leads to a particular judgement of recalcitrance. That is, to say of an experience that it is recalcitrant is to say that (1) and (3) seem to be true and (2) seems veridical but that they lead to a contradiction. They are contradictory since the theory predicts P (in virtue of (1) and the fact that E prompts assent to I in virtue of (3)), but E prompts assent to not-P (in virtue of (3)). We end up with $(P\ \&\ \sim P)$ in the web. P is held true on the basis of the theory along with its logic and the experience E, and not-P is held true on the basis of experience only. It is the contradiction which puts the scientist's beliefs in tension and forces her to make changes in the theory.

Faced with a recalcitrant experience, the scientist has many options. In Quine's picture, every statement, or chunk of the web, is revisable. The goal, when faced with a recalcitrant experience, is to revise at least one of some of these statements in order to get rid of the recalcitrance. Let's enumerate a few obvious candidates for revision:

O1: Changing the theory so that it does not yield the problematic observational conditional $(I \rightarrow P)$ anymore. (Getting rid of the eternal sentence 'Copper oxide is green').

O2: Changing the underlying logic L so that we cannot deduce the observational conditional whilst keeping the original theory. (Changing for quantum logic, say).

O3: Rejecting the experience E by pleading hallucination or by simply ignoring it. (Denying that the stuff actually turned green).

O4: Claiming that the initial conditions I were not strictly observed. (Claiming that one of the mixing tubes did not contain the appropriate substance).

Performing a revision along the lines of options (O1) to (O4) can indeed remove the recalcitrance. But Wright points out that one option is missing from that list, namely

the option to deny that E is recalcitrant in the first place by denying that $I \to P$ follows from Θ via L. Without the observational conditional, the experience E ceases to be recalcitrant. That is to say, it is at least a possible option to reject W (namely, to reject '$\Theta \vdash_L (I \to P')$'). To reject W is not to reject the underlying logic, and it is not to reject the theory. It is rather the plain denial that $I \to P$ follows from the theory via the logic L.

There is a prima facie difficulty with this. One could say: how can we reject that $I \to P$ follows from the theory via L if all the steps leading to $I \to P$ are L-accredited steps? At first blush, it does seem that rejecting a statement of the form W amounts to rejecting L, so that the option of rejecting W is not a distinct option. But this is missing the point. The fact that $I \to P$ follows from the theory via L simply is another sentence like the others. That is, it is possible that one did not follow the L-accredited steps correctly, that there was, say, a mistake in the application of the rules of L. The Quinean must acknowledge that as a possibility. As it stands, W is revisable. If it were not, that would mean that there would be, in the end, unrevisable statements in the web — that is, necessary statements, statements immune to revision. So the Quinean is clearly committed to the revisability of W — every judgement, every sentence in the web is revisable. This is the case despite the undeniable fact that it seems hard to see how one could revise such statements. A further option by which to get rid of the recalcitrance is thus simply to deny that there is recalcitrance in the first place, not by pleading hallucination or by accepting a contradiction, but rather this time via a rejection of W:

O5: Denying that E is recalcitrant via a rejection of W. That is, denying that $I \to P$ follows from the theory-with-its-logic. (Denying that the observational conditional 'if two test tubes containing respectively substance A and substance B are mixed, then we get a green stuff called copper oxide' follows from the theory-with-its-logic).

In sum, in the situation we have imagined, the scientist faces a range of options in order to accommodate the recalcitrant experience. This seems to be a fair reconstruction of what actually happens in science.[3]

After setting up this Quinean picture of science, Wright invites us to consider the following question: when is it reasonable, under this picture, to believe a statement such as W? To put the question differently, what could compel the scientist to discard option (O5)? In order to get a clear view of the situation, let us first ask how, in the Quinean methodology, the scientist is going to decide between (O1) – (O5). Presumably, Wright says, the main consideration must be the degree of further recalcitrance associated with each option. It is sound scientific methodology to choose, amongst (O1) – (O5), the option which will provoke less further recalcitrance. We just can't pick randomly, on pain of having to cope with possibly disastrous results. We must test the options in order to make the *best hypothesis*. The notion of 'best hypothesis' — the expression used by Wright — is to be understood this way: for Quine, every statement in the web is, in a sense, a hypothesis since they can all be revised. This is especially true of (O1) – (O5), since these options are put forward as possible options, *i.e.*, they are put forward tentatively and the idea is to find the *best* one amongst them. Therefore, when choosing amongst (O1) – (O5), the scientist wants to choose the *best* hypothesis between the options — this is in effect the point of the methodology. So the task of the sound scientist at this point is to assess the consequences of adopting each of (O1) to (O5), and then, on the basis of these results, to choose amongst them. To assess the consequences means, in this context, to assess the degree of recalcitrance associated with rejecting, say, the theory Θ and replacing it with another theory. Assessing the degree of recalcitrance of a given option presupposes a means to calculate this degree. How this is done precisely might vary, but there has to be such a means. In any case, to assess the degree of recalcitrance

---

[3]Although it has to be conceded that (O2) is a very rare case and (O5) has surely never happened.

involves assuming, for the sake of the argument, that the web has been changed in a particular way (one of (O1) – (O5)) and to test that hypothesis against experience. Therefore, for each of (O1) – (O5), we make new judgements of recalcitrance.

Let us now ask again, when is it reasonable to discard the option of rejecting W? In order to assess the viability of keeping W, we have to determine the degree of recalcitrance attached to it, *i.e.*, we have to evaluate the 'cost' of discarding it. If the price is too big — if abandoning W provokes a lot of recalcitrance — it's better to keep it. The cost is determined by comparing (O5) with the other options. Let us say that our scientist arrives at these results (the number represents the degree of recalcitrance):

Adopting O1: 8

Adopting O2: 25

Adopting O3: 12

Adopting O4: 4

Adopting O5: 17

The situation seems straightforward. On the basis of these results, the scientist has to adopt the fourth option on the list — she has to claim that the initial conditions described by the observational conditional were not followed correctly (since (O4) contains the least degree of recalcitrance).

But despite appearances, it is not that straightforward. Wright asks us to examine what exactly the scientist is doing when reasoning from hypotheses — when, that is, she is determining the degree of recalcitrance of each option. She is, we have just seen, making *new* judgements of recalcitrance. In other words, she is making judgements of the form (1) – (3). These new judgements are, of course, revisable statements just like the others. The question arises at this point: how do we know that option,

say, (O3) has a degree of recalcitrance of 12 is the best hypothesis? That is, how do we know that the proposition 'rejecting the experience E results in a degree 12 of recalcitrance' is the best hypothesis? It is a possibility, after all, that instead of holding such a statement true it would be better to ignore the experience(s) involved in this judgement, or to reject the W involved, or to deny that the initial conditions were observed, etc. This is because assessing the degree of recalcitrance of a given option must involve particular judgements of recalcitrance — must involve, that is, particular judgements (of the form (1) – (3) above) to the effect that adopting a given option occasions a certain number of recalcitrant experiences. In order to know which judgements of recalcitrance to keep and which to reject amongst those judgements, the scientist has to see how every judgment of recalcitrance fares in comparison with the other options. This again invites a new query as to see which hypothesis is the best, and the scientist makes new judgements of recalcitrance, and so on infinitely.

In the end, we can never know when it is reasonable to hold true statements such as W. This is because in order to assess the viability of W, we have to make judgements of recalcitrance, which are further hypotheses. But are they the best hypotheses? Since recalcitrance is always a hypothesis, we have to make sure that a particular judgement of recalcitrance is the best hypothesis to hold — otherwise believing in a judgement of recalcitrance would amount to believe in a hypothesis without rational grounds. This subtle argument invites many comments and reflections. In what follows, I shall be concerned, first, to examine one possible Quinean counter-argument, and second, to identify what exactly is undermined by the argument, if successful.

3.3 A Quinean reply

The regress in Wright's argument begins when it is observed that the new judgements of recalcitrance made in the course of deciding the best hypothesis are themselves revisable hypotheses. This means that the Quinean must then decide whether

these new judgements of recalcitrance form the best hypothesis (whether these judge-
ments of recalcitrance form the best option to hold), and regress ensues. One imme-
diate way to stop the regress is to appeal to a pragmatical maxim stipulating that
for the sake of being able to make changes in the web, the scientist must regard the
new judgements of recalcitrance as fixed and unrevisable *for the time being*.[4] To use
Neurath's metaphor, the idea here is that while the scientist is fixing her boat, she
need only worry about the immediate task at hand, *i.e.*, fixing the damage. What
happens next might be cause for further worry, but right now the scientist has to
focus exclusively on her main job. In other words, she must take the new judge-
ments of recalcitrance as unrevisable at least for a while. She does not need to check
whether they are the best hypothesis or not in order to resolve the recalcitrance that
caused the problem in the first place. According to this line of thought, the presup-
position that Wright makes — that each judgement of recalcitrance must be shown
conclusively to be the best in hypothesis in order to hold it true — is unwarranted.
The scientist can, in fact, stop the regress by accepting the new judgements of re-
calcitrance as they are. What justifies this move is a consideration of a pragmatic
nature: the scientist has to do her best to solve the problem at hand. To come back
to our example, she is justified in adopting option (O4) on the ground that even if
she does not know whether the judgements involved in choosing (O4) form the best
hypothesis, they look like they are the best (since she arrived at these results), and
that is enough for the task at hand.

This seductive argument, however, misses the point of Wright's argument. Re-
member what happens at the first stage of recalcitrance — when the first experience
E occurs that leads the scientist to make the best possible change in the web. We
have an established theory, together with its logic and its predictions. Something
goes wrong — an experience occurs and conflicts with the prediction. Now it is a

---

[4]This Quinean defence has been brought to my attention by Gerry Hough.

crucial step in Wright's argument to recognise that recalcitrance itself is a hypothetical matter (that is, that judgements of recalcitrance are themselves revisable). That means that it is possible, in principle, to solve the problem by denying that there is recalcitrance (a conflict between experience and prediction) in the first place. So it is still a matter to be decided whether there is recalcitrance or not. In other words, recalcitrance might not be the best hypothesis. To identify the best hypothesis is clearly a central methodological principle that a Quinean will accept. This is what happens when the scientist ponders on which option to choose between (O1) and (O5) — all these new judgements of recalcitrance are made to determine which option is the best hypothesis. So far, the Quinean will accept all this. But she demurs when we point out to her that she must also make sure that these new judgements of recalcitrance are the best hypothesis. This is where the regress begins in Wright's argument. The Quinean suggests that we should stop there, that there is no need to justify these new judgements for the purposes of the task at hand.

But what warrants this sudden refusal to find the best hypothesis? If we were so concerned in finding the best hypothesis in the first place (when choosing between (O1), (O2),...,(O5)), how can one justify the move that we have to take the new judgements of recalcitrance without checking whether they are the best hypothesis or not? I want to point out two things here. First, taking the new judgements of recalcitrance as they first come out is a risky business. Let us imagine that our scientist chooses (O4) since she accepts the *prima facie* result that this will occasion only a degree four of recalcitrance. She proceeds to make the changes in her web. But it turns out that these new judgements of recalcitrance concerning (O4) were not the best hypothesis. In fact, it turns out, let's assume, that instead of holding that this change would occasion a degree four of recalcitrance, it would actually occasion a degree forty-two, since upon examination it so happens that she made a mistake in the original judgement of recalcitrance or else she neglected to take some considerations

into account. The crucial point is that she took her original judgement of recalcitrance at face value, but it turns out that had she revised it, she would have realised that in fact, opting for (O4) is disastrous. Let's explain the matter differently. The scientist is now at the stage in which she is revising (O4) — upon realisation that something is wrong, after having adjusted the theory, etc. In doing so, she makes new judgements of recalcitrance, and picks the best one. Now let us imagine that when she was making these further judgements of recalcitrance, she arrived at these results:

(O1)* 56

(O2)* 42

(O3)* 75

(O4)* 45

(O5)* 99

According to these results, she ought to choose (O2)* (let's assume that this option consists of ignoring the experience causing the recalcitrance). But this is clearly a serious problem for the theory. Whereas in the very first instance of recalcitrance we had to choose amongst options containing between a degree four and a degree twenty-five of recalcitrance, we now face a situation where we have to choose between 42 and 99. The theory ends up very badly as a result of having choosing option (O4) in the first place. But this situation happens because the scientist took her preliminary judgements of recalcitrance at face value. If she had studied the consequences of adopting (O4) more carefully, she wouldn't find herself in that situation. The situation is potentially highly problematic since effecting a change in the web may occasion, in the long run, many other changes. Thus, it might become very difficult to accommodate all the new changes prompted by the realisation that she in fact occasioned a lot of recalcitrance without realising it immediately. The point here is

that taking judgements of recalcitrance without checking whether they are the best hypothesis amounts to a step in the dark — assuming, as a good Quinean should, that judgements of recalcitrance are hypotheses.

3.4   What exactly does Wright's argument establish?

These considerations bring me directly to my second point. Pushing the situation I have just described a bit further, we find that the Quinean never has a firm grasp of the notion of recalcitrance. This point was actually brought up by Wright in his paper, although he did not expand on it. The Quinean, forced to accept the *prima facie* results of judgements of recalcitrance to avoid a regress, never can be sure that an experience is recalcitrant. The only means she has to decide whether a given experience is recalcitrant is by accepting it without checking whether it is the best hypothesis or not, by stipulating that the chain of justification has to come to an end somewhere. But we have just seen that it is not sound scientific practice to proceed this way — one cannot say both that judgements of recalcitrance are hypotheses *and* that they are not subject to appraisal like other hypotheses. The Quinean can never know whether a particular experience is or is not recalcitrant. This is a serious defect. The notion of recalcitrance figures at the heart of the web of belief; it is the very notion around which the success of a theory is to be assessed. If there is no way to tell whether an experience is recalcitrant — whether it is the best hypothesis or not — then the whole foundation of Quine's system collapses. As Wright says, Quine's system can't explain the requisite notion of success (1986: 193). Indeed, it cannot explain it since success is, for a Quinean, to be explained in terms of a theory's degree of recalcitrance — yet there is no way to tell whether an experience is recalcitrant or not if we follow Quine's methodology scrupulously (if we accept the claim that every statement is revisable in the light of empirical evidence).

One way to avoid the regress is to hold judgements of recalcitrance 'true come

what may' — unrevisable. Indeed, a natural solution is to regard at least some of the ingredients of such judgements as being necessary in the sense of being unrevisable. Remember that those ingredients consist of (1) − (3):

1. $W: \Theta \vdash_L I \rightarrow P.$

2. An experience E.

3. E prompts assent to $(I \ \& \ \sim P)$.

We can't hold an *experience* to be necessary or a priori. Experiences are unconceptualised events, they just happen. So it can't be a reasonable option to hold that (2) is the necessary or a priori ingredient. Looking at (3), we find that the fact that E prompts assent to I and $\sim P$ is determined by the very content of the experience. Again, that E prompts assent to I and $\sim P$ can't be held to be necessary, since assigning a content to an experience depends on many factors — it is clearly a contingent matter. For instance, one can have an experience prompting assent to the sentence 'this is red', but it is always possible that the same experience would have prompted assent to a difference sentence given different circumstances (if the speaker knows he's hallucinating, or knows that the perceptual conditions in which he finds himself are not standard). We are left with (1), the statement of logical consequence. This amounts to saying that, given the veracity of the content assigned to the experience, under the present proposal the Quinean — or anyone — must regard statements of logical consequence as being unrevisable. To do so allows judgements of recalcitrance to have the role they are intended to play in the Quinean picture, namely to act as regulators of the web and as measure of success. If such judgements can be taken to be true in virtue of the necessity of statements such as (1), the regress stops. That is, if revising statements of the form W is not an option, then there is no need to test judgements of recalcitrance indefinitely. It is true that we have to discard the possibility of rejecting the experience involved before holding the judgement of recal-

citrance as being definitely true. But unless one is ready to espouse a naive scepticism about experience, there has to be a point where hallucination can reasonably be discarded. In other words, it is true that we can never be sure, in a sense, that the content assigned to an experience is veridical: it is always possible *in principle* to plead hallucination. But hallucination is the exception, not the rule. Once sufficient tests have been made to rule out the plausibility of rejecting the second ingredient in judgements of recalcitrance — the assent to $I \ \& \sim P$ — the necessity of W seems to provide the foundation needed to avoid the regress. This is a very tempting argument — it does fix the Quinean boat.

However, what is the notion of 'true come what may' in play here? What does it amount to? The correctness of the argument above equates logical necessity with unrevisability. This is a natural thing to do. Indeed, it seems that the problem identified by Wright with regards to the Quinean methodology is the following: the regress occurs because every statement is a hypothesis, and as such each one is *revisable*. If that is the source of the regress, then it is overwhelmingly tempting to stop it by declaring statements of the form W as being unrevisable, that is, as being 'true come what may' and thus necessary. But it is a mistake to equate logical necessity with unrevisability. This equation stems from an ambiguity in the expression 'true come what may' (an expression used by Quine repeatedly when talking about necessity). To be 'true come what may' can mean two things:

1. Whatever happens, I will hold the statement true because its truth is inviolable.

2. I am prepared to hold the statement true no matter what happens because its truth does not depend on the configuration of the world, broadly conceived.

This is a subtle distinction.[5] The first formulation does not amount to logical necessity — it merely amounts to treat the statement as having a very high degree of certainty.

---

[5] For which I am grateful to Crispin Wright.

I might be very certain of the truth of a statement without that statement being logically necessary. For instance, I am highly certain that I am right now in the postgraduate Hut situated behind Edgecliffe, but it's obviously not logically necessary. I am also highly certain that there are other minds, but despite the amazingly high probability of that being the case, it is not necessarily so. This first sense of 'true come what may', is to be equated with unrevisability, namely the idea that the statement will never be revised, no matter what happens. A logically necessary statement *is*, on the other hand, revisable. As a matter of fact, we do not often revise our judgements of logical necessity; but the possibility to do so is not discarded from the outset. We can get a proof wrong. We can make a mistake when computing a proof. Some mathematicians have made mistakes in their proofs that have been discovered years later. And we allow for that — a correct proof is logically necessary even if we allow, in principle, the possibility of revising it. Thus, the truth of our judgements of what is logically necessary is not inviolable. Locating the source of the regress to the fact that in the Quinean methodology every statement is revisable results in trying to fix the problem by treating W's as unrevisable. But doing so amounts to interpret the notion of necessity in the first sense; and that is a confusion.

What about the second sense of necessity? Does it correspond better to our intuitions about the notion? If so, how does it relate to Wright's argument against Quine? In the second sense, to be 'true come what may' means to be prepared to hold the statement true no matter what happens in the actual world, no matter how it changes.[6] The world is thus-and-so — my holding a statement as necessary means that it does not matter how the world is; it could be different and I would still hold the statement true. In fact, no matter what the configuration of the world is, I cannot foresee myself changing, on its basis, my judgement of necessity. This differs from

---

[6]One is tempted to say: To hold the statement true *in all possible worlds*, but this is a formulation I would prefer to avoid in the present thesis.

the unrevisability claim: nothing is said, in this formulation, about unrevisability as such. All that is maintained is that I cannot envisage the possibility of revising the truth-value of the statement in the light of how the empirical world is. That means two things: firstly, that even if I cannot envisage the possibility of doing so given my state of information about how the world is and could be, I can still envisage the possibility of new information coming in, requiring me to revise the judgement. Secondly, it means that I can revise the statement in the light of things other than the way the world is. For instance, as mentioned before, I could be prompted, by a mistake in a proof, to revise my judgement. But this is not a change prompted by the way the world is or could be. This seems to be a more reasonable conception of what it is to hold a statement as being logically necessary.

Now, how does this understanding of logical necessity relate to Wright's argument against Quine? If we take Wright's argument as establishing the point that a methodology in which every statement is revisable leads to a regress, we will be tempted to stop the regress by declaring some statements as unrevisable. But we can also view the argument in a different perspective: the argument might be taken to show that there are statements which must be evaluated not in the light of further recalcitrance but according to a different methodology. More precisely, one can diagnose the problem in Quine's methodology in the following way. There is a regress since all statements relate to experience. Because they all so relate to experience, the truth of each statement has to be evaluated in relation to their empirical recalcitrance, so to speak. This need to evaluate *every* statement in relation to their degree of potential recalcitrance is what causes the regress in final analysis.

If this is the right way to look at it, the source of the regress has nothing to do with revisability, but rather with the fact that all statements are revisable in the light of experience. This would call for allowing a class of statements to be evaluated outside this arena. Since the source of the regress lies in making judgements of

recalcitrance, the diagnosis is that such judgements must be evaluated outside the Quinean methodology. That is precisely what Wright concludes: 'So the reasonableness, or otherwise, of judgements of recalcitrance must be exempted from appraisal via the Quinean methodology. And that must go for the ingredients in such judgements, including statements like W' (Wright, 1986: 194). He then claims that they are to be understood in terms of proof and that this shows that we "possess some sort of concept of logical necessity" (Wright, 1986: 195). Much has been read into that claim: Ian McFetridge (1990) and Hale (1999) argue that Wright is not entitled to draw that conclusion from his argument. Hale writes:

> It is one thing to accept that we have to regard a proof as establishing a statement (such as W) in a way that is not subject to holistic appraisal, and quite another to hold that a statement thus established is *necessary*. What the argument shows, at best, is that we have to accept such statements as established, by suitable proofs, as *true*; nothing in the argument demands their acceptance as *necessarily* true. (Hale, 1999: 46)

Hale, who also defends logical necessity, uses Wright's argument in conjunction with an argument of McFetridge to show the necessity of logical necessity. According to Hale, both McFetridge's argument and Wright's are incomplete. However, Hale claims that McFetridge's argument can be made to work with a little help from Wright's argument. His strategy is convincing — it seems to convincingly show the necessity of necessity. In what follows, however, I shall pursue another approach based on the considerations just reviewed.

3.5 From proof to necessity

Now that the notion of logical necessity has been clarified and the source of the regress identified, we can proceed to show how it is possible to fill the gap in Wright's

argument. Wright claims that since some ingredients in our judgements of recalcitrance must be exempted from Quine's empirico-holistic methodology, that shows that we possess 'some sort of concept of logical necessity'. We have just seen how Hale and McFetridge dispute this inferential step. In this section, I shall defend Wright's contention and try to justify the inferential step. Wright, McFetridge, and Hale all agree on one important point: that Wright's argument establishes that statements such as W's are arrived at by *proof*. So what has been established is that there *is* a fundamental difference between a proof and an experiment. This is crucial. It is a recurrent empiricist thought that logic and mathematics are as 'empirical' as any other field and that the apparent difference in kind is merely one in degree. Thus, to put in another way, Wright's argument shows that there is a class of statements which is not empirical.[7]

An empirical statement is one whose truth is determined by, roughly speaking, 'the way things are in the world'. That means that if Wright's argument is correct, and if the interpretation given in the previous section is also correct, then the truth of statements of the form W is *not* determined by the way things are in the world. It is a small step to connect these considerations with the second interpretation of 'true come what may' given above, namely: I will hold the statement true no matter what happens because its truth does not depend on the configuration of the world, broadly conceived. A judgement about the truth of a statement of the form W is made via a proof. Such a proof, it has been suggested by Wright, lies outside the empirical arena. So when I form a judgement by way of proof, the truth of this judgement will not be affected by the configuration of the world. I will hold it true 'no matter what' — not in the sense that it will be unrevisable, but in the other sense. A revision of the statement is still possible, but such a revision will proceed through the proof

---

[7]For more on the proof / experiment distinction see Wright (1980: Part Three, esp. Chapter XVII).

methodology. Therefore, there *are* statements whose truth is not affected by the way the world is, contrary to what Quine's holism implies. Such statements are logically necessary.

3.6 Conclusion

This Chapter was concerned to study Wright's argument against the holistic picture of belief advocated by Quine. This picture of our relation to the world is in effect a manifestation of Error-theoretic Irrealism, since Quine claims it has no need for true statements of the form $\Box(\Gamma \vdash_L \Theta)$. Wright's argument purports to refute Error-theoretic Irrealism by showing that the Quinean methodology suffers from an important problem: it can't explain what it is for a theory, or a change in a theory, to be better than another one. But as such, it does not establish without doubt the need for a notion of logical necessity, as pointed out by Hale and McFetridge. However, once we reflect on what Wright's argument establishes and on what we mean by a judgement of logical necessity, it is possible to bridge the gap in Wright's argument. In the next Chapter we study another form of scepticism about logical necessity, namely the substitutional account of logical consequence.

CHAPTER 4

ERROR-THEORETIC IRREALISM: LOGICAL

CONSEQUENCE

4.1 Introduction

Error-theoretic Irrealism about logical necessity is the view that the notion is, in some sense, disposable. We have seen that Quine takes holism to be a picture of our cognitive relation to the world which entails Error-theoretic Irrealism. However, our assessment of Quinean holism led to the conclusion that this epistemological picture does on the contrary require a notion of logical necessity. It is needed both to make sense of the logical interconnections between statements in the web of belief and to make sense of the very idea of recalcitrance. However, there is another standpoint from which Quine espouses Error-theoretic Irrealism. Quine, qua logician, owes us an account of logical truth and logical consequence. That is, he must explain in which sense these truths are *logical*, as opposed to being non-logical truths. What, then, according to Quine distinguishes logical truths from the other truths? This is the question with which we shall be concerned in this Chapter.

This is a crucial question for Quine the logician, even though his epistemological theory seems to imply that nothing distinguishes logical truths from other kinds of truths. There is indeed a prima facie tension between Quine's epistemological views and his concern to explain what it is that makes logical truths 'special'. I shall not attempt to diagnose the source of this tension. Rather, I will examine Quine's account of logical truth and logical consequence as it stands on its own. In any case, his account — to be faithful to Error-theoretic Irrealism — must avoid any notion of logical necessity. His task is to make sense of the strong intuition that logical truths are somehow 'different' from the other truths without invoking the most plausible reason to believe the intuition — without, that is, saying that these truths are special because they are logically necessary.

The Chapter is structured as follows. The next three sections discuss and examine Quine's account of logical truth. Strawson's objection to the account is scrutinised and I show how a Quinean can answer that objection. In section 4.4, I put forward

an argument to the effect that the notion of meaning has to figure at the core of the account, contrary to what Quine holds. The remainder of the Chapter (sections 4.5–4.12) looks into Quine's substitutional account of logical consequence, in effect an extension of his account of logical truth. Tarski's account of logical consequence is also examined in relation to Quine's. The main strategy in these sections is, firstly, to show that Quine is committed to the Tarskian account; and then to argue that the Tarskian account has to be interpreted in terms of representational semantics, which is an irreducibly modal account of logical consequence. The main purpose of these sections, then, is to show that Quine (indeed, everyone) is committed to a modal account of logical consequence — a non-modal account simply does not work.

4.2   Quine's account of logical truth

Quine (1951) divides analytic truths into two classes. The first class consists in the logical truths proper, truths such as 'no married man is unmarried'. The second class consists of statements reducible to logical truths proper by putting synonyms for synonyms, such as 'no bachelor is a married man'. In section 1 to 4 of 'Two Dogmas' Quine focuses on the second class. The special characteristic of the first class is that their logical form is valid. That is, putting appropriate variables in place of the names and predicates, we get a valid schema. For instance, 'no unmarried man is a married man' can be analysed like this (taking 'G' for 'is a man' and 'F' for 'is married'):

$$\sim \exists x ((Gx) \,\&\, (Fx \,\&\, \sim Fx))$$

By contrast, if we perform the same operation on 'no bachelor is a married man', we get (taking 'H' for 'is a bachelor'):

$$\sim \exists x ((Gx) \,\&\, (Fx \,\&\, Hx))$$

This is not a valid schema. The sentence 'no bachelor is a married man' will be interpretable as yielding a valid schema only if we can show that in this particular case, 'bachelor' and 'unmarried man' are synonymous expressions, so that we can transform the sentence into 'no unmarried man is a married man' by way of putting synonyms for synonyms. This latter sentence exhibits a valid logical form — it is a logical truth proper. The reducibility of the second class to the first class clearly depends on the notion of synonymy. As we have seen in Chapter One, Quine holds that synonymy can't be used satisfactorily to perform the reduction, and this is his main argument to the effect that sentences such as 'no bachelor is a married man' are true simpliciter, on the same level as 'the cat is on the mat'.

Whilst in the first Chapter we were concerned with the status of the reductive class (statements reducible to logical truths), the focus of this Chapter will be on the logical truths. In Chapter One, we saw that Quine argued in 'Two Dogmas' that the so-called analytic statements of the second class were not analytic — that is, true by virtue of meaning. Were this kind of statement reducible to the first class, it would show that they are, in some sense, logical. This division of analytic truths into two mutually exclusive classes enables Quine to attack the standard notion of analyticity without endangering logical truths: if logic is not analytic in the sense of 'true by virtue of meaning', that means that his attack on analyticity does not affect the status of the logical truths. But then, what is so special about the statements of the first class? In which sense are they analytic? It is undeniable that there is something special about 'if Socrates is a man, then Socrates is a man', in contrast with 'Socrates is Plato's teacher'. Clearly Quine also believes that logical truths proper have a special status. After all, he is the author of five introductory books on logic and the philosophy of logic.[1] His account of logical truth and of logical

---

[1] *Mathematical Logic* (1940), *Elementary Logic* (1966a), *Methods of Logic* (1950), *Philosophy of Logic* (1986), *Set theory and Its Logic* (1969b). He is of course the author of many logical papers, some of which assembled together in his *Selected Logic Papers* (1995b).

consequence is therefore crucial with regard to Error-theoretic Irrealism, since if he can explain the special status of logic without invoking logical necessity or, for that matter, any intensional notion, he will have succeeded in steering between the Scylla of intensionality and the Charybdis of the assimilation of logical truths into 'regular' truths.

Let's now introduce Quine's account of logical truth. In a nutshell, a statement is logically true, in Quine's sense, if it is not merely true but remains true under all uniform substitutions of its component expressions other than the logical constants.[2] Let us call this a *substitutional account* of logical truth. This account was first suggested by Bolzano in his *Wissenschaftslehre* (1837).[3] To get a clear idea of the structure of the account, let's consider first an example of propositional logic. Take for instance:

$$\text{If Socrates is moral, then Socrates is mortal} \tag{4.1}$$

That sentence is clearly a logical truth, if anything is. The substitutional account explains its logicality via the ideas of valid schema and substitution. The logical form of an expression is obtained by replacing its non-logical components with variable terms. Taking, as we are analysing (4.1) under propositional logic, 'Socrates is mortal' as an atomic expression, the logical form of (4.1) is:

$$\lceil \phi \to \phi \rceil \tag{4.2}$$

The central idea of the substitutional account is that (4.1) is a logical truth since its

---

[2]Quine (1951: 22). Quine actually uses the word 're-interpretations' instead of 'substitutions'. But since the former has semantical connotations, I prefer to use 'substitution', a term he does use elsewhere.

[3]Partly translated into English under the title *Theory of Science* (1972). Bolzano's account is framed in terms of propositions and ideas, contrary to Quine who uses statements and expressions. But the central idea of *substitution* to account for logical truth and logical consequence is clearly Bolzano's. It is easy to recast Bolzano's account in terms of sentences and expressions.

logical form will produce only truths when the propositional variables are substituted with appropriate expressions of the object language. That is, whatever one substitutes for 'Socrates is mortal' in (4.1), the result will be true since (4.1) is an instance of (4.2), and the latter is a valid schema. Quine understands validity of schemata in terms of metatheorems: a schema is valid if it is a metatheorem. In the case of a tautology of propositional logic such as (4.2), its corresponding metatheorem — $\vdash \lceil \phi \rightarrow \phi \rceil$ — says that whatever you may select, if ϕ then ϕ. Of course, there are constraints on substitutions — more on that later. An example of quantificational logic is:

$$\text{No unmarried man is a married man} \tag{4.3}$$

This is a logical truth at the subsentential level. That is, it is the internal 'structure' of the sentence, so to speak, which is logically true, rather than a combination of sentences. To explain the logicality of (4.3), we need a mode of logical form which 'goes into' the logical form of sentences. Quantificational logic is just such a mode. Taking for instance 'G' for 'is a man' and 'F' for 'is married', we get the logical form:

$$\sim \exists x ((Gx) \ \& \ (Fx \ \& \ \sim Fx)) \tag{4.4}$$

The idea here is the same. Whatever permissible substitution is made for 'F' and 'G', the result will be a true statement.

So one of the key ideas of the substitutional account of logical truth is that of logical form. The following distinction will help to put the account into a wider perspective. There are broadly two kinds of account of logical truth, those that make central appeal to logical form, called *interpretational* accounts, and those appealing to modality, called *representational* accounts. A representational account is one that uses modal notions such as necessity and impossibility. It often exploits possible world semantics. It will, for instance, say that (4.3) is logically true since it is true in

all possible worlds. It is clearly in Quine's interest to put forward an interpretational account.[4] The contrast between interpretational and representational accounts will be discussed further in section 4.7.

As mentioned, there are constraints on permissible substitutions. The first constraint is that of sameness of grammatical category. For instance, we cannot perform a substitution on (1) so that we get:

$$\text{If is a horse is mortal for, then is a horse is mortal for,} \tag{4.5}$$

by substituting 'is a horse is mortal for' for 'Socrates is mortal'. Such a substitution does not result in a true statement. The reason for that is clear enough: 'is a horse is mortal for' simply isn't a well-formed statement. The same applies to logical truths of quantificational logic. Names must be substituted by names, and predicates by predicates. It is important to see that this constraint does not rely on intensional notions. In effect, the constraint of grammatical category says that you have to respect the grammar of the terms you substitute. For the sake of simplicity, let's say that we're dealing with these three categories: statements ('ϕ'), names ('x'), predicates ('F'). What we have to assume here is a language with formation rules and a stock of expressions to serve as names and predicates, as well as a stock of logical operators (more on the latter below). Formation rules and the grammatical partitioning of expressions between statements, names and predicates are part of the syntactical part of the language. We need know nothing about either the meaning or the reference of the names and predicates of the language when performing a given substitution. All we have to know is the grammatical category of both the expressions we're substituting and the substituted expressions. Therefore, the constraint of grammatical category does not presuppose the notion of meaning.

---

[4]Tarski's account, of which more in section 4.7, is also an interpretational account.

The second constraint on substitution is that of uniformity. If a substituted expression occurs more than once in the original statement, we must substitute all occurrences of that expression with the same substituting expression. For instance, since the component expression 'Socrates is mortal' appears twice in 'if Socrates is mortal, then Socrates is mortal', one has to substitute both occurrences of 'Socrates is mortal' with the same expression. Otherwise, 'if Socrates is mortal, then Aristotle is the King of France' would count as a permissible transformation of 'if Socrates is mortal, then Socrates is mortal' and therefore the latter would be disqualified as a logical truth. The same consideration applies to the logical truths of quantificational logic: like names must be replaced by like names, and like predicates by like predicates. Otherwise, the account will break down. Again, the constraint of uniformity seemingly has nothing to do with meaning. All it says is that the substitutions have to be performed in a uniform manner. That is to say, if you perform a substitution in a given expression, you have to make sure you make the same substitution if the component expression you substituted occurs more than once in the original statement. This is a mechanical task, involving simply a recognition of grammatical category and identity of expressions — apparently no meaning in play here.[5] We will see in the next section how Strawson disputes this claim.

There is a third very important feature of the substitutional account, one that has been much discussed in recent literature. Quine's account presupposes a set $\mathfrak{F}$ of fixed terms, that is, terms that we cannot substitute. For instance, when considering 'if Socrates is mortal then Socrates is mortal', only the non-logical terminology ('Socrates', 'mortal') is substitutable. We cannot substitute the words 'if' and 'then', otherwise we could arrive at a falsity. It is of course intended that $\mathfrak{F}$ will consist in the logical terminology — these are the terms that are usually held fixed. Given a

---

[5]The issue is complicated by so-called category mistakes, for instance 'if 6 is funny then 6 is funny'. It seems that in these cases meaning must be taken into account, even if grammar is respected. But we will ignore this complication in what follows.

set of logical constants, we want to say that the substitutions must occur among the non-logical terminology of the language. This way of proceeding has been implicit since the beginning of the section. We took for granted a correct division of terms into logical and non-logical. For instance, '&' and '→' were implicitly part of $\mathfrak{S}$. We did not substitute them. But it is beyond question that under the substitutional account, what counts as a logical truth will depend on what one takes $\mathfrak{S}$ to be.[6] How to distinguish the logical terms from the non-logical terms is much debated. It is clear that in order to obtain the 'right' set of logical truths (whatever that means), we have to include appropriate terms in $\mathfrak{S}$. We have an idea of what the logical terms are: we want '∼', '→', '&'. We don't want: 'Socrates', 'is blue'. However, there is no consensus on how to divide logical terms from non-logical terms, and no consensus on whether such a division is desirable or even feasible. Worse, there are borderline cases, such as identity and set membership.

For his part, Quine takes for granted the obvious class of logical terms, and he discusses at length whether or not we should accept the borderline cases (identity and set membership) amongst the logical terms. He offers no method to distinguish precisely between logical and non-logical terminology.[7] For the time being, we will assume that there is some kind of *pragmatically* justified distinction. That is, I will take the class of logical terms to consists in what are generally considered to be logical terms in logic books. The list includes some or all of these: conjunction, negation, the conditional, disjunction, equivalence, the quantifiers. This alone shows that we know intuitively what a logical term is, since we know what the list must include. Concerning the borderline cases, we will have more to say on them in sections 8

---

[6]Bolzano was aware of the fact that a different set $\mathfrak{S}$ produces different logical truths: 'It is obvious that the validity of one and the same proposition must turn out differently depending on whether we look on this or that, just a single one or several ideas within it as variable' (Bolzano, 1837: §147, p.191).

[7]Tarski offers a criterion in a posthumously published paper (1986). Shapiro (1998: 145) holds that the fact that there is no objective way of classifying terms into logical and non-logical is enough to discredit an exclusively formal approach.

and 10. For now, we shall assume that even if we don't know where to draw the distinction, it does not necessarily mean that there is no such distinction.[8]

4.3 Strawson's argument against the substitutional account

Strawson (1957) discusses Quine's account of logical truth. His paper centres on the criteria for suitable substitutions. Quine holds that in order to make the substitutions, no intensional notions are required. Substitutions have to meet the requirements of sameness of grammatical category and of uniformity, and observance of these does not, at first blush, involve use of intensional notions. In his paper, Strawson argues that mere observance of these two requirements, unsupplemented by the notion meaning, is not enough. To make the account work (*i.e.*, to make it produce the logical truths we want it to) the notion of 'meaning the same as' is required. If Strawson is right, it follows that Quine either has to accept a notion of meaning or else has to abandon his account. Since, arguably, the substitutional account is the most economical one in terms of sparseness of intensional and metaphysical notions, it is hard to see how Quine could abandon it. Strawson's criticism thus represents a serious threat to Quine.

Strawson's main point is that Quine's account cannot rule out that an ambiguous expression such as 'he is sick' threatens to turn 'if Socrates is mortal, then Socrates is mortal' into a non-logical truth. The reason is this. The expression 'he is sick' has, according to Strawson, two interpretations (1957: 16). One the one hand, it may mean that a male person is *physically* ill. On the other hand, it may mean that a male person is *mentally* ill. On this view of the matter, substituting uniformly 'he is sick' for 'Socrates is mortal' in 'if Socrates is mortal, then Socrates is mortal' will result in the statement 'if he is sick, then he is sick', which may be either true

---

[8]See Shapiro (1998: 145) and Tarski (1935: 418-420) for a discussion of this topic.

or false (it is a contingent statement) since it is possible that one is physically ill without being mentally ill, and vice-versa. There is no a priori assurance that the substitution will result in a true statement. So it seems that merely observing the uniformity constraint is not enough. Replacing like expressions by like expressions can reduce what seems to be an uncontroversial logical truth into a 'mere' truth, thus devoiding it of its apparently 'special' status that the account is supposed to explain. Strawson goes on to discuss a number of possible constraints formulated with the aim of ruling out the permissibility of substituting the two occurrences of 'Socrates is mortal' with 'he is sick' understood in one occurrence in the 'physical' sense and understood in the other occurrence in the 'mental' sense.

The upshot of his enquiry is that such a criterion is possible only if we make use of the notion of meaning (Strawson, 1957: 22-23). That is to say, in order to block the unwanted substitutions, we have to stipulate that the substituting expressions must have the same sense, or the same meaning, or, in the case of statements, express the same proposition. This stipulation has the effect of ruling out our problematic case since in order to produce a 'contingent' statement from 'if Socrates is mortal, then Socrates is mortal' by substituting 'he is sick' for both occurences of 'Socrates is mortal', the two occurrences of 'he is sick' must have a different meaning. If they have a different meaning, the substitution is not permitted under the new constraints, thus preserving the logicality of 'if Socrates is mortal, then Socrates is mortal'.

Strawson is right in that the existence of ambiguity in a language, as such, forces us to take into account the meaning of the expressions to 'save' the status of 'if Socrates is mortal, then Socrates is mortal' as a logical truth under the substitutional account. Let's state at the outset that we do not wish here to dispute the soundness or validity of Strawson's reasoning. The problem with this objection to Quine's account is that it does not identify its target correctly. Ambiguity exists only in natural languages. According to Quine, the phenomenon of ambiguity results from the fact

that sometimes a word has two different applications (1960: 129-30). For instance, the predicate 'is sick' has two extensions. It applies both to a mental condition and to a physical condition. Because of that, failure of communication sometimes results and a remedy is in order — in this case, paraphrase (Quine 1960: 157-61). It is clear that Quine, when discussing logical truth and logic in general, is concerned with a regimented language in which ambiguity is removed through suitable paraphrase. The Quinean remedy to the ambiguity of 'sick' would be to paraphrase it by distinguishing between 'physically ill' and 'mentally ill'. The idea is to eliminate the ambiguous word 'sick' in the regimented language and replace it with both 'mentally ill' and 'physically ill', thus removing the ambiguity. The ambiguity is detected and eliminated without invoking intensional notions. In other words, the first step is to realise that the word has two extensions, and this step involves no intensional notions. The second step is to defuse the ambiguity by paraphrasing the ambiguous expression into two distinct expressions, each of which is to be applied to one of the two extensions detected previously. Again, no intensional notions are needed to perform this step.

This move clearly answers Strawson's worry in a neat way since in the regimented language there simply isn't the expression 'he is sick'. Quine wants, in effect, to study logic as it is manifested in the sciences and also in a properly regimented everyday language. If everyday language is filled with ambiguous expressions, we must regiment it first and then study its logical properties. Since ambiguity is a defect of natural languages and that the languages he is concerned with do not contain ambiguous expressions, it follows that his account of logical truth is not subjected to Strawson's criticism.

It is clear that there is no winner nor loser in this debate. We are rather faced with a clash between two conceptions of philosophy. Strawson, at the time, was an advocate of 'ordinary language philosophy', whereas Quine is a direct descendant of logical positivism whose main preoccupations were with science and artificial languages.

Quine, it is true, is more interested in natural languages than his predecessors. But we must not lose sight of the fact that Quine's natural languages must be regimented before being given a full-fledged philosophical analysis. For all of Quine's criticism of logical positivism, it is safe to assert that he shares the interests of his teachers in this respect. But despite these differences with Strawson, it can nonetheless be maintained that the clash in question is a methodological one and that Strawson and Quine both share an interest in more or less the same philosophical problems. For Quine, paraphrasing and, more generally, departures from ordinary language are "aids to understanding the referential work of language and clarifying our conceptual scheme" (1960: 158). Compare this with what Strawson says in *Individuals*: "my aim is to exhibit some general and structural features of the conceptual scheme in terms of which we think about particular things" (1959: 15). Viewed from this perspective, it appears that the main difference with Strawson is, after all, one of method rather than one of overall philosophical interests. Indeed, *Individuals* and *Word and Object* can be seen as two different attempts to come to grip with the same question — that of how language relates to the world.

We have seen in this section how Quine's regimented approach to the study of our conceptual scheme contains the means to answer the Strawsonian objection to his account of logical truth. I now wish to present another argument against the Quinean account which is in the same spirit as Strawson's but different from it.

4.4 The need for meaning

The argument levelled by Strawson against the substitutional account was to the effect that it makes an implicit use of the notion of meaning, contrary to what Quine claims. We have just seen how a Quinean has the resources to answer Strawson's objection. In this section, I present a different argument purporting to show that the notion of meaning is required. Strawson's target was one of the requirements for

suitable substitutions (the constraint of uniformity). Here, however, we will focus on the set $\mathfrak{F}$ of fixed terms. Remember that under the substitutional account of logical truth, the logical constants form a fixed set $\mathfrak{F}$ of non-substitutable expressions. A statement's logical truth is thus a function of this set of logical constants. Tarski (1935: 418-20) and Bolzano (1837: 191) were acutely aware of this.[9] Quine, on the other hand, does not go into elaborate discussion about $\mathfrak{F}$ — in fact, he is strikingly silent about the whole topic, seemingly taking for granted an appropriate choice of $\mathfrak{F}$. His only suggestion about the nature of the logical constants is, as far as I can see, that they are basic to all discourse (see for instance (Quine, 1940: 2)).[10] He would certainly be sympathetic to the approach we adopted, that is, to consider as members of $\mathfrak{F}$ those terms which are considered logical constants in logic textbooks, and to discuss the borderline cases individually.

It is clear that under the substitutional account, what will be considered a logical truth depends crucially on the choice of $\mathfrak{F}$. Were proper names or predicates members of $\mathfrak{F}$, we would have different logical truths. I would like to go further and suggest that what is to count as a logical truth depends not only on the choice of $\mathfrak{F}$, but also, and importantly, on the *interpretation* given to the members of $\mathfrak{F}$ — namely, the logical constants. The reason is this. Let's assume that the fixed set includes 'not' and 'or'. Now let's consider 'Socrates is mortal or Socrates is not mortal'. Under the Quinean account, this is a logical truth (of propositional logic) since whatever statement you correctly substitute for 'Socrates is mortal', it will produce a true statement. But how do we know that? That is, how do we know that all permissible substitutions will produce true statements? It is important to note that the substitutional account does not answer that legitimate question. To get a sense of the importance of that

---

[9]So are, of course, all contemporary writers on the topic, such as Etchemendy (1990), Shapiro (1998), Sher (1996), and others.

[10]In his (1986: 62), Quine also talks about the universality of logical constants. This universality is what makes it tempting to accept identity in the class of logical terms. But he does not elaborate on this, nor does he recommend adopting universality as a criterion of logicality.

question, consider the statement 'every even integer greater than 3 is the sum of two prime numbers'. To say that 'Socrates is mortal or Socrates is not mortal' is a logical truth under the substitutional account, the statement 'every even integer greater than 3 is the sum of two prime numbers or every even integer greater than 3 is not the sum of two prime numbers' must be true as well. But such a statement might not be taken to be true by the proponents of intuitionistic logic, therefore threatening the status of 'Socrates is mortal or Socrates is not mortal' as a logical truth, and, more generally, threatening the validity of the schema $\lceil \phi \text{ v} \sim \phi \rceil$. Quine, a proponent of classical logic, will of course consider all instances of the schema as logical truths — that is, he thinks (to use his terminology) that $\lceil \phi \text{ v} \sim \phi \rceil$ is a metatheorem. Proponents of, say, intuitionistic logic do not consider $\lceil \phi \text{ v} \sim \phi \rceil$ as being a metatheorem. How should one explain this divergence?

I take the example to show that what is regarded as being a logical truth under the substitutional account depends on how we interpret the set of fixed terms consisting in the logical constants. In other words, it depends on the meaning assigned to the logical constants. The proponent of intuitionistic logic will be able to find an appropriate substitution to render the supposedly logical truth 'Socrates is mortal or Socrates is not mortal' into a statement she will not assert since, according to her, it is plainly not the case that every statement exemplifying $\lceil \phi \text{ v} \sim \phi \rceil$ is assertible. And that is so because for her, negation and disjunction are respectively assigned different meanings from those assigned by the classicist. This, in any case, is what I shall defend in the following paragraphs.

It is an undeniable fact that classical logic and intuitionistic logic do not share the same stock of logical truths. Under the substitutional account, this is made clear by the fact that the two logics will arrive at a different set of logical truths — the extension of their accounts differ. Let us take two logicians, one classicist, Carl, and one intuitionist, Igor. Let us assume that both subscribe to the substitutional account

of logical truths. Remember that Quine's account of logical truth merely consists in 'cashing out' the idea that a sentence is a logical truth if it is true and remains true under any substitution of its non-logical components. Carl and Igor likewise proceed to cash out this idea, by laying down the constraints of sameness of grammatical category and of uniformity and by determining what the set $\mathfrak{I}$ of fixed terms consists in: for instance, they might agree to include '→', '&', '∼'. So far, their description of the theory is exactly the same on both sides: when Carl and Igor present their account of logical truth, there is absolutely no way to distinguish between them. Let us say that they both present it for first-order logic, and that they choose, as members of $\mathfrak{I}$, negation, implication, and disjunction. It is clear and incontestable that there is no way to distinguish between the two accounts as described so far. In both cases, the theory will say that a logical truth is a sentence which is true and remains true under all substitutions of their non-logical components. Both theories of logical truth, at this point, will be identical, and make use of absolutely no intensional notions (they are staunch Quineans).

However, we know that despite the descriptive identity of their account, Carl and Igor will not agree with each other on the extension of logical truths — that is, on which statements pass the substitutional test and are declared as being logical truths. They will argue over the logicality of, say, 'Socrates is mortal or Socrates is not mortal'. For Carl, it *is* a logical truth; for Igor, it is not. Nothing in their theory of logical truth can account for this difference in extension. This is the crucial point: it is impossible, by simply 'reading off' their theory, to predict the divergence. The primary goal of a theory of logical truth is, presumably, to generate the 'correct' extension of logical truths. However, the account as described by Carl and Igor offers no guarantee that it will produce their favourite stock of logical truths. Something is missing from the account — an additional feature that will ensure that each logician arrives at his favourite set of logical truths. And it is this extra feature over which

Carl and Igor will argue.

The missing component in the account is the interpretation of the set $\mathfrak{F}$ of fixed terms. This, or so it seems to me, is the only way to explain the extensional divergence between Carl and Igor. The substitutional account has to be supplemented with an explanation of how the logical constants are to be interpreted — and this is something that is lacking in Quine's description of his theory. One way to understand the divergence between classical logic and intuitionistic logic is by appealing to the introduction and elimination rules assigned to the logical constants — the so-called logical schemata. Carl will say that negation is governed, partly, by double-negation elimination. Igor will balk at this. Differences of this kind are at the basis of the divergent output of substitutional theories. To be fair to Quine, it should be mentioned that he does explain validity in terms of schemata, conceived as giving the forms of logically true sentences. For instance, he makes it explicit that since $\lceil \phi$ v $\sim \phi \rceil$ is a valid schema, it follows that every statement exhibiting its form is logically true. But for Quine, schemata are mere pedagogical aids.[11] That is why he does not use them in *Mathematical Logic*. But even in that book, when he talks about, say, $\lceil \phi$ v $\sim \phi \rceil$ (not a schema, remember, but a metatheorem in *ML*), what makes it a valid metatheorem? In the system of *Mathematical Logic*, there is no answer to that question. The natural answer would be that the 'truth' of $\lceil \phi$ v $\sim \phi \rceil$ hinges on the meaning of 'v'. But this answer is not available to Quine.

In view of this, another way of fleshing out the disagreement between Carl and Igor is to point out that a crucial feature of the substitutional account is often missing in Quine's description, namely the reference to valid schemata. For Carl, $\lceil \phi$ v $\sim \phi \rceil$ is a valid schema, whereas for Igor it is not. The point here is the same: the reason

---

[11] *Methods of Logic* proceeds in terms of schemata. *Mathematical Logic*, however, avoids their use. He says, in the 1981 preface of *Mathematical Logic*, that 'schemata and even open sentences are technical aids'. He does not use them in *Mathematical Logic* because he fears that readers will consider sentence letters — such as 'p', 'q' — as taking *propositions* as values, whereas for Quine we can only talk about *sentences* (see Quine (1940: iv-v)).

why one accepts the validity of the schema and the other resists it has to do with the meaning attached to the sign '∼'. There seems to be no way to escape the following conclusion: the meaning of the set of fixed terms has to be stipulated in the account itself. It has to be built in. There is therefore an irreducible notion of meaning at play in Quine's account of logical truth.

This is where the deep issue with the substitutional account lies. It takes for granted an underlying interpretation of the logical constants. That is to say, it takes for granted the interpretation — the meaning — of the set of fixed terms. *Given* the logic, *given* the interpretation, the substitutional account certainly can be made to work to a certain extent, once problems such as ambiguity are resolved in a Quinean spirit. But let's not forget that the substitutional account requires these assumptions, and because of that, it does, ultimately, rest on the notion of meaning. The statement 'Socrates is mortal or Socrates is not mortal' is logically true for the classicist by virtue of the meaning of negation and disjunction. That is, it is logically true because of the introduction and elimination rules assigned to negation and disjunction. Once you are equipped with that interpretation of negation and disjunction, every substitution (performed under the constraints mentioned above) will result in a true statement — even the substitutions that result in the statement 'every even integer greater than 3 is the sum of two prime numbers or every even integer greater than 3 is not the sum of two prime numbers'. One who prefers to give an intuitionistic interpretation of negation will not consider 'Socrates is mortal or Socrates is not mortal' as a logical truth if he subscribes to the substitutional account.

One of the net effects of this argument against the substitutional account is that Quine's division between two classes of analytic statements collapses. Both kind of analytic truths, the logical truths and the statements reducible to logical truths, are after all dependent on the notion of meaning. If, as I have argued in this section, the substitutional account is committed to the view that logical truths are analytic

in the standard sense (that is, true by virtue of meaning), it follows that Quine's attack on analyticity is also an attack on his own account of logical truths. It does not show that the notion of meaning in play has to be a Platonistic one, but it shows at the very least that Quine, if he wants to retain his account of logical truth, must presuppose *a* notion of meaning. And that in turn might be problematic.

The issue of how Quine should handle this problem goes beyond what can be discussed in this thesis. But what I take it to show is that Quine is committed to the view that logical truths are analytic: so he is, or so it seems, committed to the view that there is a notion of necessity in play in his account, and this for two reasons. Firstly, Quine has it that analyticity and necessity go hand in hand, as seen in Chapter One. Secondly, even if we acknowledge a conceptual distinction between analyticity and necessity, it remains that if a sentence is analytic in the sense of being true by virtue of meaning, it follows that it is necessary. That is, if a true sentence owes its truth to the meaning of the logical constants appearing in it, that sentence *has* to be true.

These considerations make it plain that Quine's account of logical truths does *not* do without intensional notions. More to the point here, it seems that Quine's account presupposes a notion of necessity. If that is correct, then it is not the case that Quine 'can do without necessity' — that he can be an Error-theoretic Irrealist about necessity. Perhaps the only live option for a Quinean who wishes to 'save' the substitutional account is to acknowledge the requirement of analyticity and to cash it out in a conventional manner — to 'go conventionalist'. It is well-known that Quine himself has no sympathy for conventionalism.[12] Another option for Quine is to drop the notion of logical truth altogether: but this is not an option he entertains at any point.

---

[12]At the very least, he has no sympathy for conventionalism of the Carnapian kind. See in particular Quine (1936b and 1963).

4.5   Quine's account of logical consequence

Now that we have explained and discussed Quine's account of logical truth, we have to explain its natural extension, namely his account of logical consequence. It is clear that the topic of this thesis is related in an important way to logical *consequence* (see Introduction). What concerns us here, in effect, is the status of logical inference, not of logical truth. However, an account of logical truth can easily be adapted to yield an account of logical consequence, so that all the criticisms raised about the former apply to the latter. In fact, as we will see, logical truth can be seen as a limiting case of logical consequence. In this section, I will explain how Quine's account of logical truth extends to an account of logical consequence, and in the following sections I will raise some problems related to this new account, problems which are better discussed in relation to logical consequence than in relation to logical truth.

It is often heard that Quine does not consider logical consequence to be the main topic of logic, that for him the study of logic is the almost exclusive study of logical truths. But I think this is an oversight. Quine himself repeatedly stresses the primary importance of inference in logic. For instance, in his 'Introduction' to *Methods of Logic*, he writes: "The chief importance of logic lies in implication, which, therefore, *will be the main theme of this book*. Techniques are wanted for showing, given two statements, that the one implies the other; herein lies logical deduction" (1950: xvi, my italics). Other passages display a similar view.[13]

Let's now describe Quine's substitutional account of logical consequence. As I have said, this theory is to be viewed as an extension of his theory of logical truth. Both theories make central use of the ideas of substitution and logical form. To begin with, let us state some intuitions about the concept under study. Logical consequence is a relation holding between a non-empty set of statements (the premises) and another

---

[13]See also (1950: 33; 1940: 7). In *Philosophy of Logic*, Quine does seem to restrict his attention to logical truth, but this is simply for the sake of convenience (see (1986: 48-49)).

statement (the conclusion). The idea is that a statement 'logically follows' from a non-empty set of statements if it is impossible for the premises to be true and the conclusion to be false. This way of wording logical consequence, however, is very un-Quinean since it makes use of modal notions — it is equivalent to saying that given the premises, the conclusion *necessarily* follows.[14] The objective of Quine's account is to arrive at the 'correct' extension of logical consequences — we want any account to get the extension right — without using intensional notions. The case here is entirely parallel to logical truth: in that case also, Quine wanted to give a purely extensional account. He was only concerned with getting the extension right and avoiding intensional notions. For Quine, the 'correct' extension consists of those logical consequences that can be arrived at via classical logic.

The substitutional account of logical consequence is designed to meet these desiderata. According to this theory, a statement follows from a non-empty set of other statements provided that whenever permissible substitutions are performed on the non-logical terms, the conclusion will be true provided the premises are.[15] The main difference between the account of logical consequence and that of logical truth is that in the former case, there has to be talk of *relation* between sentences, whereas in the latter case we are only focusing on the status of a single sentence. In an account of logical consequence, we want to know the status of a conditional consisting of a set of sentences as antecedent and a single sentence as consequent. If it shown that any permissible substitution of the components in the conditional results in a true conditional, we will have shown that this conditional displays a valid logical form. For instance, consider the following inferential judgement: 'All men are mortal and Socrates is a man. Therefore, Socrates is mortal.' We clearly want to say that the conclusion follows from the premises, and any decent account of logical consequence

---

[14]The equivalence is seen by the fact that '$\sim \Diamond(\Gamma \,\&\, \sim A)$' is intuitively equivalent to '$\Box(\Gamma \rightarrow A)$'.

[15]Notice that there is no intensional idiom used in that definition.

must explain that. As we did with logical truths, the first step is to exhibit the logical form of the argument:

$$((\forall x(Fx \to Gx)) \,\&\, Fa) \vdash Ga \tag{4.6}$$

Now it is clear that whatever we put for 'F', 'G', and 'a' the statement formed at the right-hand side of the turnstile will be true provided that the statements forming the antecedent are also true.

4.6 The requirement of persistence

In this section and the following ones, I will discuss the main criticisms brought against the substitutional account of logical consequence as well as the counter-arguments offered by Quine. There are basically four lines of objection against the Quinean account of logical consequence. They have to do with:

- The dependence of the account on the richness of the language.

- Worries about the restriction on first-order logic.

- The status of identity — logical or non-logical constant?

- Worries about the size of the intended domain.

I will take up these objections and worries one by one. My strategy will be to show that Quine is forced, on the basis of these arguments, to embrace a stronger, less austere account of logical consequence to meet the objections and worries raised. As the objections are discussed, it will become clear that he is committed to a version of the Tarskian account, a version irreducibly appealing to our modal intuitions. If this is right, my argument will show that Quine cannot embrace a purely extensional account of logical consequence.

The criticism raised in the present section reveals the main drawback of talking in terms of substitution, namely that the account will be hostage to the expressive

resources of the language. Let us consider the following language. It contains only, as extra-logical terms, the names 'Lincoln' and 'Clinton' and the predicates 'is President' and 'is American'. Consider now the sentence:

$$\text{Lincoln is President} \tag{4.7}$$

The problem is immediate: under the language specified, every permissible substitution will result in a true statement. Therefore, (4.7) will be declared a logical truth. But clearly, (4.7) is not a logical truth.[16] The same difficulty appears when we consider logical consequences such as:

$$\text{Clinton is American} \models \text{Lincoln is President} \tag{4.8}$$

The substitutional account of logical consequence will declare 'Lincoln is President' to be a logical consequence of 'Clinton is American' in this language, since for any permissible substitution on these two statements, if the former is true so will the latter. Obviously, however, it is not an intuitively valid argument. Passing the substitution test might therefore not be a sufficient condition for logical truth and logical consequence. The substitutional account *overgenerates* in this case; that is, it counts as valid what should not be declared valid. What these examples show is that the correctness of the account is dependent on the richness of the language examined.

It has to be admitted that the language just defined is not a very interesting language. But the point is general. Consider, for instance, the vaguely defined semi-scientific language used to talk about the world and to make sense of it. Presumably, the universe treated by this language extends what can be expressed by it. We don't have a word for every object in the universe. If we consider any logical consequence — for instance, the one in the previous section — the account declares that 'Socrates

---

[16]Of course, Lincoln isn't a President right now, same for Clinton. But we ignore this detail.

is mortal' is a logical consequence of the premises, and this is what we want. It will tell us (via the logical form exemplified by (4.6)) that whenever permissible substitutions are performed on the non-logical terms, the conclusion will be true provided the premises are. However, the account does not take into consideration the expandability of language. That is, it offers no guarantee that were a new term to make its appearance, the result of operating a substitution with this new term will not render invalid a substitution performed in accordance with (4.6). It appears, in contrast with the examples discussed in the previous paragraph, that the substitutional account offers no guarantee that it can *undergenerate*; that is, there is no guarantee that the substitutional account will count as valid what should be considered valid.

In sum, as Tarski put it, the substitutional account will work "only if the designations of all possible objects occurred in the language in question. This assumption, however, is fictious and can never be realized" (1935: 416). Of course, *we know* that any instance of (4.6) is valid, and *we know* that the addition of a new term will not invalidate it under the substitutional account (given an appropriate choice of $\Im$). This problem, in effect, is just a different side of the same difficulty encountered with the limited language of the previous paragraph. In the case of a restricted language, the account clearly overgenerates, and in the case of our more inclusive language, it has the potential to undergenerate. The central point here is that Quine's theory cannot account for a central feature of logical consequence, namely that an instance of correct logical consequence *persists* through expansion of the language — logical consequence should not be hostage to the expressive power of the language under study.

The upshot is that a correct account has to meet what Etchemendy (1990: 30-31) calls *the requirement of persistence*. It can be stated thus:

(RP) If Θ is said to be a logical consequence of Γ under an account of logical consequence, that theory must guarantee that it will continue to be so through

expansions of the language.

The substitutional account does not meet this requirement, since it is confined to the actual expressive resources of the language. The reasoning here is that Quine's account offers no guarantee that a given logical consequence will remain valid (under the account) if we expand the language. This is because expanding the language means that there will be new substitutions — and the account offers no guarantee that these new substitutions will not transform what was considered a valid argument into an invalid one.

At this point one might raise the following objection. It is true that Quine's account does not meet the persistence requirement, but that needn't worry Quine since he isn't after a characterisation telling a story about what *could* happen were the language to expand. He is only concerned, the objection goes, with circumscribing the actual extension of logical consequence, that is, with what are in fact the valid inferences expressible in the language. Viewed in this light, it seems that the substitutional account has got the extension right, given an appropriate choice of fixed terms. Even if Quine's account is not completely faithful to our intuitive notion of logical consequence (since the requirement of persistence is part of this intuition), it does not matter because he thinks that what matters in a definition of this sort is to get the extension right, not to be one hundred percent faithful to our vague intuitions. Indeed, it is a virtue, from Quine's point of view, that the theory does not contain any notion that can be interpreted as a modality, such as what *could* happen were the language to expand. At this point Quine can appeal to his holism and say that the substitutional account is true as far as it goes, that it is the best hypothesis we've got, etc. If an unfortunate expansion of the language occurred, a Quinean could say, we would perhaps have to revise our theory, but we can be pretty confident it won't happen.

However, a Quinean cannot rest content with this reply. A definition of logical consequence would hardly be worthy of its name if it captured only those consequences that actually occur in the language under study. It is obvious (given the intended domain, *i.e.*, the 'external world') that there are other logical consequences that are inexpressible so far and even more that will never be expressed. An account of logical consequence failing to meet the requirement of persistence does more than merely conflict with our intuitions: it's not an account of logical consequence proper, but merely an 'anthropological' account — in the sense that it tells us what follows from what for us given the actual history of our language. Besides, the Quinean needn't worry much about the modality in play here since it is a minimal kind of modality: to recognise the possibility of the expansion of language is not an endorsement of the necessary / contingent distinction as it is understood by, say, the logical positivists. To say that the language might expand involves the same sense of modality as to say that it might rain tomorrow. It is certainly not this sense of possibility that Quine's philosophy of logic is trying to avoid.

The Quinean, therefore, must confront the challenge brought about by the requirement of persistence. Two options are possible for the Quinean: she can either convince us that the substitutional account does meet the requirement despite what Tarski and Etchemendy claim, or she can modify the account so that it satisfies the requirement. In his *Philosophy of Logic*, Quine addresses this matter, albeit not in these terms. However, to get a clear view of the issue, we have to make a detour through Tarski's account of logical consequence.

4.7   Tarski's account of logical consequence

Tarski's account, developed in his much-discussed (1935), was designed specifically to meet the requirement of persistence. What we need in order to overcome this

difficulty is an account similar[17] to the substitutional one but which uses, instead of the notion of substitution, the notions of *satisfaction, sequence,* and *model.* For the purposes at hand, we need only an intuitive characterisation of these notions. Let us take the open sentence 'x is the capital of y'. The *sequence* of objects ⟨Paris, France⟩ *satisfies* this open sentence, whereas the sequence ⟨Tokyo, Spain⟩ does not. The notion of a model can be characterised thus. Keeping with the same example, we say that the sequence ⟨London, Great Britain⟩ is a *model* of the sentence 'Paris is the capital of France' via the open sentence 'x is the capital of y' and this same sequence (⟨London, Great Britain⟩) satisfies the open sentence. It is important to notice that sequences here have a dual role: they satisfy open sentences and they model closed sentences. With this intuitive terminology at hand, we can define Tarskian logical consequence thus:

(**TLC**) Θ is a logical consequence of Γ if and only if every model of the sentences in Γ is also a model of Θ.

The idea here is similar to the substitutional account in one important respect: in both cases, a sentence follows from another if and only if the whole argument exhibits a valid logical form.

One could say that the Tarskian theory represents our intuitions about logical consequence more faithfully, since it respects the requirement of persistence (**RP**). Indeed, the notions of sequence and satisfaction, in contrast to those of substitution and truth, generate an account of logical consequence that is independent, as Tarski put it, of the richness in concepts of the language being investigated. The notion of substitution is hostage to the expressive power of the language: we can only substitute with expressions that are available. The notion of satisfaction, on the other hand, deals directly with objects in the universe of discourse, independently of whether

---

[17]I say 'similar' since both accounts are what Hanson (1997: 366-67) calls 'formal accounts of logical consequence' and what Etchemendy (1990: 51) refers to as 'interpretational semantics'.

these objects are named or not in the language under study. It deals with the whole domain of discourse, in contrast with the notion of substitution which deals only with the objects of the domain that are provided with a name in the language.

Tarski's account, then, meets the requirement of persistence. It shares with Quine's account the emphasis on logical form, and it does not appear to involve modalities.[18] These facts should make the account attractive to a Quinean. Nevertheless, Quine makes it clear that he favours his own account over the Tarskian model-theoretic one (1986: 53-56). This is because the model-theoretic account is committed to an ontology of sets, whereas the substitutional account requires only a minimal ontology composed of expressions in the language. The Tarskian account is indeed committed to an ontology of sets since the sequences are, in effect, ordered sets, and the account quantifies over these sequences. This commitment to sets was to be expected. If we want to go beyond the names and predicates that are expressible in the language in order to take into consideration the objects and the properties that are not named in this language, we will inevitably have to make reference to the objects of the domain as such, not only the named objects. It is this unavoidable reference to objects, via the concept of a sequence, that commits one to sets when putting forward the model-theoretic account. In the Quinean account, *expressions* are substituted; in the Tarskian one, sequences of *objects* are satisfied.

At this point one could raise the following objection.[19] Why is Quine making such a big fuss over the commitment to sets? After all, he embraces set theory and is the author of an introduction on the topic (1969b). Maybe the young Quine had some reservations about set theory and its ontological commitments, but the post-fifties Quine is certainly happy with set theory and its commitment to sets. Indeed, Quine is a Platonist concerning mathematics. But this objection misses the point. Quine is

---

[18]It relies only, as Hanson (1997: 370-71) puts it, on truth *simpliciter*. More precisely, in the Tarskian account, on satisfaction *simpliciter*.

[19]Graham Priest raised this objection in a seminar.

trying to define logical consequence. It is a logical notion, and as such it should be definable without using non-logical notions. If, when defining logical consequence, we are forced to acknowledge a commitment to sets, we are not doing logic anymore but rather set theory. There is nothing wrong with set theory; but it's not logic. More on the relation between set theory and logic in section 4.9.

Another important feature of Tarski's account as defined here has to do with domains. To make that feature clear, I will now explain a crucial distinction that has been alluded to in previous sections. According to Etchemendy (1990), there are two ways to understand model theory. One is to say that the models are an interpretation of the language. Under this understanding of models, the domain stays fixed and what varies is the meaning (the interpretation) of the terms. When we say, under this understanding of models, that every model of a non-empty set of premises is also a model of the conclusion, what is said is that every true interpretation of the non-logical terms in the premises will result in a true interpretation of the conclusion. For example, under the interpretation that 'man' means 'whale', 'mortal' means 'mammal' and 'Socrates' means 'Jamie', the models thus formed from (4.6) will result in a valid logical consequence. Notice here that nothing has been mentioned about variation of domain. Under this understanding of the notion of model, the domain stays fixed while it is the interpretation assigned to the non-logical terms that changes. In effect, to use a less technical language, what is happening here is that we are testing for validity by enquiring whether the premises and the conclusion of the argument are true and remain true no matter which interpretation (extension) is assigned to its non-logical terms. If it remains true under all interpretation, then it corresponds to a valid logical form. If it corresponds to a valid logical form, the interpretation assigned to the non-logical terms can vary at will. According to Etchemendy, this understanding of models is the one implicit in Tarski's celebrated paper on logical consequence. It is also consistent with the exposition of Tarski's account as given

here, since in this exposition no talk of variation of domain is mentioned, nor does it need to be. This understanding of what models are is called the *interpretational* view of models.

It is clear that Quine's account is also a brand of interpretational account. This is seen by the fact that the substitutional account, just like Tarski's, makes crucial use of the notion of logical form. It is also seen, more conspicuously, by the fact that Quine's account is also concerned with what happens if the interpretation of the non-logical terms is changed. In Quine's case, of course, we are not talking about varying meanings but only expressions. However, the idea is the same: an argument is valid if it is true and remains true under all *substitutions* of its non-logical terms — if it remains true no matter which interpretation we assign to the non-logical terms.

The other understanding of the notion of model is the representational one. In this case, the models represent possible configurations of the world. This is the familiar notion of model currently used in contemporary model theory. To put it simply, when we are saying that 'All men are mortal; Socrates is a man; therefore Socrates is mortal' is a valid argument under this understanding of model, we are in effect saying that, for instance, in every possible world where Socrates is a man and all men are mortal, Socrates is mortal. We are not varying interpretations here, neither are we using the notion of logical form, but, rather, are considering what happens in a particular possibility: we are considering a possible world. This kind of account is uncongenial for a Quinean, since it is clearly modal. We will have much more to say about representational accounts of logical consequence later on in this Chapter.

To recapitulate, there are two broads kinds of account of logical consequence: interpretational and representational accounts. Tarski (1935) uses the notion of model, though not the contemporary notion. It is best understood as an interpretational account. Quine's account is also an interpretational account, but it does not use models. In contrast, the current, model-theoretic account of logical consequence is a

representational account. Quine, in a nutshell, has no choice but to avoid any kind of representational account. Concerning the Tarskian account as exposed in the (1935) paper, Quine wants to show that his account has the same explanatory power as Tarski's but is more economical since it contains no ontological commitments. We will now address this matter.

4.8   Meeting the requirement of persistence

Let us now return to the question that prompted this detour to the Tarskian account. What is Quine's attitude towards the fact that his preferred account does not meet the requirement of persistence? He is well aware of the potential discrepancy between his theory and Tarski's: "there is no assurance that each class of objects in our universe corresponds as extension to some open sentence constructible from the words of our language" (1950: 95). He then goes on to explain that the question of whether we should choose open sentences (required by the substitutional account) as suitable interpretations of schemata or classes of objects (required by the Tarskian account) "depends on our choice of universe, and also on how rich a vocabulary we assume there to be at our disposal" (*ibid*). This means that if the vocabulary is rich enough, there is no need to appeal to classes of objects and we save on ontology. So it is not as if Quine first put forward his account and then realised that it does not meet the requirement of persistence. He was always aware that his account has the potential to undergenerate. The main reason to repudiate the Tarskian account had to do, as early as in his (1950), with ontology: "(i)t is sound policy to confine the assumption of philosophically contested entities, in particular abstract entities such as classes, to those portions of theory which need them" (1950: 95, footnote 1). The maxim of ontological economy is in play here. Don't make appeal to Tarski's account, Quine warns us, unless you have to. The argument studied in this section is that according to Quine, we do not need Tarski's account to make sense of logical

truth and logical consequence — the substitutional theory suffices.

Quine recognises the limits of his account when he says that it assumes "a fully interpreted object language with no option left open as to range of variables" (1986: 52). But, he points out, the substitutional account will work in such a fully interpreted language. Elementary number theory is expressible in such a language: first-order logic (**L1**), with or without identity.[20] In this particular language and universe, all the logical consequences that are captured by the model-theoretic account are also captured by the substitutional account. That is, the two definitions coincide in extension when restricting our attention to first-order logic with or without identity. In other words, the following principle of equivalence holds for **L1**:

(**EQ**) (Θ is a Quinean logical consequence of Γ) $\equiv$ (Θ is a Tarskian logical consequence of Γ).

The left to right conditional is provable via the downward Löwenheim — Skolem theorem:[21]

(**DLS**) If a schema in **L1** is satisfied by a model, then it is satisfied by a model whose domain is at most denumerably infinite.

Now, how do we get from (**DLS**) to our desired result? Let's notice first that the antecedent of (**DLS**) makes no reference to a domain. The antecedent, more precisely, is about satisfied schema independently of the domain. The consequent is about satisfied schema in a particular kind of domain, that is, those that are at most denumerably infinite. The positive integers form such a domain. So (**DLS**), in effect, reduces Tarskian validity in any domain to Tarskian validity in a denumerably infinite or finite domain. In view of the fact that the positive integers form a domain of the

---

[20]Quine, of course, prefers the version without identity, where it is paraphrased in terms of predicates.

[21]See *e.g.* Quine (1986: 54) and Shapiro (1991: 80).

latter sort and that for such a domain the substitutional account is applicable (since all the objects of the domain are expressible in the language), we get from (**DLS**):

(**T1**) If a schema is satisfied by a model, then it becomes true under some substitution of sentences of elementary theory.

By contraposing (**T1**), we obtain:

(**T2**) If a schema is false under all substitutions of sentences of elementary number theory, then it is satisfied by no model.

If, instead of talking of a schema in (**T2**) we talk about its negation, we get our left-to-right conditional:

(**T3**) if a schema is true under all substitutions of sentences of elementary number theory, then it is satisfied by every model.

The right-to-left conditional of (**EQ**) is proved via Gödel's completeness theorem:

(**COM**) If a schema of **L1** is satisfied by every model, it can be proved deductively.

The idea, presumably, is that if Θ is a Tarskian logical consequence of Γ, then we can prove it. If we can prove it, that means that the language has the expressive resources to perform the proof, so that we will be able to carry out the substitution test.

(**EQ**) is an important principle for Quine. The fact that it holds for first-order logic shows that if we restrict our attention to it, we can do without modalities and without an ontology of sets. Saving on ontology and avoiding intensional notions are of course two important methodological principles for Quine. No wonder that he says that the provability of (**EQ**) for first-order logic is 'a remarkable fact' (1986: 53, 55). Quine's objective, we have seen, is to provide an account of logical truth that is exclusively extensional and that avoids any dubious ontological commitments. By putting forward a brand of interpretational account, he claims to have achieved the

first objective. By resisting the Tarskian interpretational account and by sticking to his substitutional account, he claims to have achieved the second. In the next section, I want to examine whether Quine succeeds in resisting the Tarskian account.

4.9 Incompleteness and second-order logic

Remember that for (**EQ**) to hold, we have to restrict our attention to languages which are formalised in first-order logic. The immediate reason for this restriction has to do with the fact that we cannot prove completeness and the Löwenheim-Skolem theorem for higher-order languages. This means that we cannot prove the crucial theorem (**EQ**) mentioned earlier:

(**EQ**) (Θ is a Quinean logical consequence of Γ) $\equiv$ (Θ is a Tarskian logical consequence of Γ).

An obvious counter-example to (**EQ**) for second-order logic are the Gödel sentences. Such sentences are Tarskian logical consequences when we interpret mathematics under second-order logic: in every model satisfying the axioms and the rules of inference of the theory, there will be a Gödel sentence which will also be satisfied (under the assumption that the axiomatic set theory is consistent). But, since the Quinean account will be applicable only to languages for which we have a complete proof-procedure, it follows that in this account there is no way to capture the fact that the Gödel sentences are consequences of the theory. In other words, when focusing on second-order languages and when equipped only with Quine's account, we don't get the right extension of logical truths and logical consequences. When applied to second-order logic, Quine's account can only capture consequences that are syntactically deducible. Gödel sentences are not syntactically deducible. This is because in second-order logic, truth exceeds provability — assuming classical logic and a realist framework.

The substitutional account, then, does not work when applied to elementary number theory formalised in second-order logic — contrary to the Tarskian account. But, as is well-known, this needn't worry Quine since for him, second-order logic is not logic: it is set theory in sheep's clothing.[22] The main reason for viewing second-order logic as set theory in disguise is that in second-order logic we quantify over predicate letters — in second-order logic, expressions such as '$\forall F \exists G \forall x (Gx \lor \sim Gx)$' are well-formed sentences. Once this is admitted, it is tempting to see quantification over predicate letters as ranging over attributes. But for Quine, attributes are not well-defined objects, as is the case with propositions and meanings; since we have no clear identity-criteria for attributes, we can't accept them in our ontology.[23] Now the natural option to avoid this objection would be to view quantification over predicate letters as ranging over sets, which are well-defined entities. But this is, again, a deplorable move to make according to Quine. Predicate letters, to paraphrase Quine,[24] do not *name* unspecified *sets*, they rather *stand in place* of *names* of unspecified sets. For instance, when writing 'Fx', we do not thereby mean that 'x belongs to F' (such talk commits one to an ontology of sets in second-order logic), but we are rather talking about an unspecified name of a set. The 'F' denotes neither a specific set nor an arbitrary set: it is a name of an arbitrary set. And one can't quantify over schematic predicate-letter, one can only quantify over objects in the universe of discourse (similarly, we don't normally quantify over the 'p''s and 'q''s of propositional logic). For that reason, if one admits sets in one's ontology (as one should, since they are well-defined entities), one has to quantify over them by using the terminology of set-theory, in which sets are objects over which we can quantify. But if one does that, one is not doing logic as such, but set theory. This is because one is dealing with a

---

[22]See Quine (1986: Chapter 5). For a full discussion of this issue, see Shapiro (1991).

[23]See Chapter One, section 1.5 for a discussion on identity-criteria.

[24]Quine (1986: 66).

specified domain of discourse, and logic is supposedly 'topic-independent'.[25]

This argument to the effect that second-order logic is set theory in disguise has been widely discussed. Here is not the place to discuss whether Quine's argument is successful — a separate thesis would be needed in order to do that.[26] But we can assess the consequences of accepting and of resisting his argument in the context of examining Quine's account of logical consequence. If we are to accept Quine's argument, there is no proper second-order logic. There is only first-order logic and set theory. When we are doing set theory, we use the axioms of set theory coupled with first-order logic, and we get the results we want.

Now if we are to resist Quine's argument instead of accepting it, the situation is straightforward. If second-order logic *is* logic, then it is clear that the substitutional account does not work. There will be sentences that logically follow from the axioms of second-order logic but not captured by Quine's account. That means the substitutional account does not get the extension right in this case. In sum, the justifiability of restricting the account to first-order logic depends on this controversial argument about the status of second-order logic. If one accepts Quine's argument, the substitutional account survives. But if one does not accept the argument, Quine's account collapses. Suffice to say, for the present purposes, that Quine's argument against second-order logic is highly controversial.[27]

---

[25]This exposition of Quine's argument against second-order logic is somewhat sketchy. But for the purposes at hand, it is enough since I am not going to dispute the argument.

[26]The interested reader should be directed to Shapiro's *Foundations without Foundationalism: A Case for Second-Order Logic*.

[27]See Shapiro (1991) and Boolos (1975; 1984) for a defence of second-order logic against Quine's arguments.

4.10 Objections relating to size of the domain

I would now like to consider another objection that has been put forward against the substitutional account. The objection purports to show that there is an intuitively invalid inference that comes out valid according to the substitutional account. That is, it purports to show that Quine's account hasn't got the extension right. Let's consider the following inference:

$$(\exists x)(\exists y)(x \neq y) \models (\exists x)(\exists y)(\exists z)((x \neq y) \mathrel{\&} (y \neq z) \mathrel{\&} (x \neq z)) \qquad (4.9)$$

This sentence says, in effect, that if there are two things then it follows logically that there are three things. On the assumption that '=' is a logical term and that the universe contains three or more objects, (4.9) will be declared valid by Quine's and Tarski's accounts since it is true and all the expressions in it are logical — there is nothing to substitute, nothing to model. In fact, this argument will be considered valid by any interpretational account.

The problem, of course, is that intuitively the conclusion does not follow from the premise in (4.9). From the fact that there are two things, it does not follow logically that there are three things. It is after all a possibility that there are only two things. The problem faced by the interpretational account in this case is that what is considered valid will vary in function of the size of the domain. In a universe containing exactly two objects, the account will consider the argument invalid. In a universe containing three or more objects, it will consider the argument valid. Thus, the account makes validity dependent on such contingent facts as the size of the universe.

One way to escape this objection is by appealing to an axiom of infinity. If such an axiom is assumed, then it will logically follow that if there are two objects then there are three objects. Quine seems to regard the axiom of infinity as merely a hypothesis,

to be assumed only when needed to prove theorems requiring it.[28] He would certainly not regard it as a logical axiom, and even less as a necessary one. It just isn't a logical 'fact' that there are infinitely many objects — logic doesn't involve itself in the way the universe is. It might be a mathematical fact that there are infinitely many objects, but even that is disputed. The appeal to the axiom of infinity is not, simply put, an available answer for Quine to the problem raised by (9).

Another line of response available to a Quinean would be to point out that in the language under consideration and its domain (first-order logic used to formalise elementary number theory, in which the domain is denumerably infinite), there is no problem to account for the fact that under the interpretational account, (4.9) is valid. The universe under consideration is infinite — given that it is, it does follow that if there are two objects then there are three objects. The idea here is to assume the infinity of the domain and then to say that under this assumption, it is correct to say that the argument depicted above is a valid one. The assumption about the size of the domain does not involve a necessary claim such as postulating the axiom of infinity. It can be seen either as a working hypothesis or as a fact. The success of this ploy, however, comes at a price. Remember that intuitively, we don't want (4.9) to be valid. Assuming infinity, or working with an infinite domain, has the result that the conclusion in (4.9) follows but at the price of being counter-intuitive. In this sense, the resulting definition of logical consequence will fail to capture what we mean by logical consequence. What is logically valid shouldn't depend on the size of the domain — however, it seems that the Quinean has to put forward a view that forces

---

[28]For instance: "The axiom has been decried on the ground that the question whether there are infinitely many individuals is a question rather of physics or metaphysics than of mathematics, and that it is incongruous to make arithmetic depend upon it. Whitehead and Russell were apologetic about the axiom, as they were about the axiom of choice; they entered both of them as explicit hypotheses in the theorems that required them, just as I did with most comprehension assumptions..." (1969b: 280-281). Also, from a logic paper in the thirties: "In particular the multiplicative axiom and the axiom of infinity are as requisite to (Quine's system) as they are to Zermelo's system; I follow *P.M.* (Principia Mathematica), however, *in suppressing these as postulates and requiring their statement rather as explicit hypotheses wherever necessary*" (1936a: 85, my italics).

the recognition that what follows from what depends on the size of the domain. As a last resort, Quine could retort that he is not concerned with getting our intuitions right, only with getting the extension right. Given the infinity of the domain, he could say, (4.9) is valid. The other immediate problem with this, of course, is that such a definition depends on non-logical truths such as the size of the domain. If sets are not allowed in the definition on the basis that they are not logical, why should contingent facts be allowed? Their inclusion in such a theory is even more contentious than sets. On this view, logical form is not sufficient to assess the validity of a logical consequence: it has to be supplemented with extra premises. This seems unfortunate.

Quine's most effective response to the problem generated by (4.9) is to avoid it by treating identity as a non-logical term. If '=' is not a logical term, then there will be substitutions (interpretations) making the argument false, thus stripping it of its validity. Again, the effectiveness of this strategy comes at a price. Quine himself is aware that identity does seem to belong more to logic than to mathematics (1986: 64). He nevertheless proceeds to reduce statements involving identity to statements involving only truth-functions and quantification (1986: 63-64). The price to pay is to regard what seems to be a logical constant as a non-logical term. Quine's arguments that identity is non-logical are highly disputed, and it is fair to say that most logicians nowadays regard it as a logical constant. It is enough to point out, here, that Quine's way out of (4.9) depends again, as is the case with second-order logic, on a contentious, highly disputed argument.

In sum, we see that Quine faces counter-examples that have to do with the size of the domain. This is related to his clause according to which the language must be interpretable under the domain of the natural numbers, an infinite domain. Quine is explicit about this. His proof of (EQ) will succeed only if that clause is accepted. What we are doing here, amongst other things, is assessing the price of this restriction. We are about to see how that clause takes central stage.

4.11   Size of the domain: Part II

One might be tempted to think that Quine, by appealing to his argument against the logicality of identity and by stipulating that the domain is infinite, gets out of trouble (even if his argument is contentious). But matters are not so simple. Shapiro points our attention to a sentence of first-order logic not involving identity which again puts the interpretational account in difficulty. Consider, Shapiro (2002) tells us, the following sentence:

$$(\forall x)(\forall y)(\forall z)((Rxy \ \& \ Ryz) \to Rxz) \ \& \ (\forall x)(\exists y)(Rxy) \models (\exists x)(Rxx) \qquad (4.10)$$

The sole non-logical term in this sentence is the binary predicate 'R'. This sentence, in effect, says that if an arbitrary relation is transitive and that every object stands in this relation with another object, then there is an object for which this relation is reflexive. Now in a finite universe, every substitution performed on (4.10) is true, rendering it into a logical truth according to the account. Think for instance of the relation '$>$'. In a universe containing, say, only the numbers from 1 to 5, the relation '$>$' is transitive but not every object stands in this relation (the number 5 is not greater than any other number in this universe). So the antecedent is false, making (4.10) true. A similar point applies to other transitive relations. But if the domain is infinite such as the domain of the natural numbers, such a sentence is not a logical truth under the substitutional account (or, for that matter, any interpretational account) since when we interpret 'Rxy' as '$x<y$', it comes out false. That is, the relation '$<$' is transitive and every object stands in this relation (for any number, there is a smaller one), but no number is smaller than itself. Here again, we have a formula whose logicality under the Quinean account depends on the size of the universe. And this time, Quine cannot rely on the status of identity.

Just as (4.9) is intuitively not a valid logical consequence, (4.10) is clearly not

either. The point is the same as before: an account that would render it as a logical truth would be an incorrect account. The only resource available to Quine, again, is to rely on his condition that the account works for the language and, even more crucially, for the domain specified. If the universe is interpretable in terms of the denumerably infinite series of the natural numbers, then sentences like (4.10) will not be considered logical truths, thus 'saving' the extension. This is because the natural numbers form an infinite domain. Again, Quine has to fall back upon considerations to do with the size of the intended domain.

So, to effectively solve the problems caused by (4.9) and (4.10), the Quinean (and Tarskian) must make assumptions about the size of the universe — she must say that the account works in infinite domains, and that these are the interesting domains. But — apart from the reservations expressed previously — another potential counter-example threatens the workability of this ploy. Let's grant to Quine that the universe under consideration is infinite and that there is a way to accommodate (4.9) and (4.10). There still remains the issue about infinitary inferences, such as the so-called ω-rule. This rule says that if a formula A is true of every natural number, then the formula 'for every number n, $A(n)$' is true. The natural numbers form an infinite domain. So here we have arguments consisting of an infinite set of premises and a conclusion. Intuitively, we would think that the ω-rule is valid: if something is true of every natural number, then we should be able to infer validly the proposition 'for every number n, $A(n)$'. But the compactness of first-order classical logic blocks this intuitively valid logical consequence. A consequence relation is said to be compact if "any consequence of an infinite set of propositions is a consequence of some finite subset of them" (Read, 1995: 43). The problem is that the premises of an application of the ω-rule cannot be reduced to a *finite* subset. Since it cannot, it follows that in first-order classical logic the ω-rule cannot be considered to be a valid rule of inference. The validity of the ω-rule is a contentious issue, highly disputed. But intuitively, it is

hard to see why someone would like to repudiate the validity of infinitary inferences. If something is true of all objects, isn't it then logically correct to infer that every object has that property?[29] Why reject infinitary inferences? The only reason to reject it seems to be that it is not compact. But this is not, in itself, a good justification. We see here how Quine faces a potential counter-example even if we grant to him the restriction on the size of the domain.

4.12 Varying domains

Now that we have looked at all these objections to the interpretational account and their replies, it is time to take stock and look at the situation from a broader perspective. To avoid problems having to do with second-order logic (for instance, Gödel sentences), Quine has to restrict the applicability of his account to first-order logic. We have seen that he offers an argument to the effect that second-order logic is not logic. So the first condition for the workability of his account is:

(C1) Restriction on first-order logic.

We have also seen that Quine, to meet objections involving arguments containing identity (see (4.9)), has to convince us that identity is not part of the set $\mathfrak{F}$ of fixed terms. Doing so enables the account to declare as invalid arguments which would otherwise be declared, counter-intuitively, valid. So Quine has to:

(C2) Treat identity as an extra-logical term.

In view of statements such as (4.10), which are intuitively invalid but declared valid under the substitutional account in a finite domain, Quine has to stipulate that the account is applicable only in an infinite universe, such as the natural numbers:

---

[29]I totally share Read's attitude on this issue. See, again, his (1995: 45). Etchemendy (1990) also holds this rule to be valid. Hanson (1997: 396, see also footnote 42) discusses it in relation to our topic, and so does Shapiro (2002: 6), but they do not take a stand on this issue in their papers. Tarski discusses the issue in his classic paper on logical consequence (1935: 410-412) and has more to say about it in his paper on truth (1933a: 258-261). See also his (1933b: 294-295).

(C3) The universe must be interpretable under the domain of the natural numbers.

We have seen, in the previous sections, that the arguments in support of restricting the applicability of the account via (C1) and (C2) are far from conclusive. I will now discuss in greater detail the third requirement.

Let's grant Quine our adherence to (C1) and (C2). Quine still faces his most serious problems with condition (C3). The latter is necessary for Quine, we have seen, in order to avoid objections relating to the size of the universe. If the universe under study is re-interpretable in terms of the natural numbers, it ensures that the substitutional account gets the extension right. In fact, this requirement amounts to a requirement concerning the size of the universe. Quine's account will work only if the universe is interpretable in terms of the theory of the natural numbers. The natural numbers form a denumerably infinite universe. So in order for Quine's account to get the extension right, we have to assume the infinity of the universe. This is problematic. For one thing, although it seems probable that the universe is infinite, it is not proved to be true. For another thing, even if it is infinite, it could have been finite. Ordinarily, we would like to say that what counts as a logical truth should not depend on the size of the universe — should not, that is, depend on contingent matters. Quine, of course, is only concerned with getting the *extension* right, he is not concerned with getting our *intuitions* right. However, a point similar to one made earlier applies here. If an account of logical consequence does not capture our intuitions about what follows from what, in what sense is it an account of logical consequence?

We know that Quine does not like to talk in terms of 'models' and 'domains': for him, there is only one domain, and that is the (assumed to be infinite) universe. In this case, we have seen, his account works (provided that (C1) and (C2) are also satisfied). But most people would like to say that there are such things as valid and invalid logical consequences in finite domains, and that any correct account should

get the extension right in these cases too. I can't see any reason why talk of 'finite domains' and 'models' should be forbidden. Quine's account cannot do the job here.

I think the above discussion of (C1) – (C3) shows that the substitutional account of logical truth and logical consequence is ultimately unsatisfying. There is no satisfying justification for restricting the applicability of the account in the way that (C1) – (C3) forces us to do. Faced with this situation, and assuming she accepts to drop these clauses, how can a Quinean minimise the ontological and modal commitments? To drop clause (C1), we have seen, commits one to a different interpretational account, the one exposed by Tarski in his (1935). But how does Tarski's account fare if we drop clauses (C2) and (C3)? This is a tricky question. To answer it, we have to consider in what sense the Tarskian account is a model-theoretic one. (C2) and (C3), in effect, relate to the size of the universe. We know that for Quine, there is only one 'domain' — the actual universe. If it is infinite then his account steers clear of problems such as (4.10). As we have seen in section 4.7, in Tarski's definition of logical consequence there is no mention of domains. He only talks about 'arbitrary sequences of objects' satisfying sentential functions. One has the impression that there is only one domain — the universe. He is however explicit that any definition of logical consequence has to be relativised to some particular language. Nevertheless, it does seem that we can have many different languages all 'talking' about the same domain (the same universe). Acknowledging the possibility of different languages does not commit one to a multitude of 'domains'. Etchemendy, we have seen, seems to think that Tarski's account is merely a variation of the substitutional one, and that both accounts are what he calls 'interpretational', that is, with varying only the interpretation of the non-logical terms, not the domains. If that is the case, then Tarski's account, as it is defined in his (1935), is no more satisfying than Quine's since it is also hostage to clauses (C2) and (C3).

The crucial point is that in order to free ourselves from these unjustified clauses, we have to be able to vary the domain — we have to be able to talk not just about the actual universe, but also other domains. Let's grant, despite the fact that it is controversial, that Tarski's account is merely a refined version of Quine's. Tarski's account, as it stands, does not allow us to vary the domain. The only thing allowed is to vary how we interpret the non-logical terms — that is, how the extension of the non-logical constants is assigned. These variations, under our assumed interpretation of Tarski, all occur within one and the same domain. Since this is the case, then Tarski's account is hostage to the size of the universe. If it contains only one object and the identity symbol is logical, it will consider as logically false sentences such as '$\exists x \exists y (\sim (x = y))$'. Otherwise, it will consider such sentences logically true. But it seems that the logicality of this sentence — of any sentence — has to be independent of the size of the universe. A similar problem occurs with (4.10). In an infinite domain, the Tarskian account will take (4.10) to be non-logical. But in a finite domain, (4.10) would count as a logical truth. Again, the objection is that whether (4.10) is a logical truth or not has to be a matter independent of the size of the universe.

The upshot is that in order to meet these criticisms, a correct account has to allow for variation of domains. That is, it has to be able to make sense of our intuition that a logical truth is true no matter how the world is configured. We want to say that a logical truth keeps its logical status no matter the size of the universe — no matter if there is one, zero, or infinitely many objects in it. This is why we have to be able to talk of a sentence being true in every *domain*, where a domain is a description of how the world is. To do this is to do what Etchemendy calls 'representational semantics'. A representational semantics is a semantics constructed in terms of contemporary model theory. This kind of semantics is very different from the kind of semantics Tarski developed in his (1935) — remember that we assume here that Tarski's definition of logical truth is interpretational. In Tarski's

account, what we vary are the interpretations assigned to the non-logical terms. In contemporary model theory, however, what we vary are the models, where a model is to be understood as a description of a possible configuration of the world. For instance, a sentence is logically true, in model theory, if it is true in all models — in all possible worlds.

In representational semantics, we don't encounter problems with either (4.10) or truths involving identity. (4.10) is clearly not a logical truth under representational semantics since there is a model where it is false, namely a model with infinitely many objects. Same with '$\exists x \exists y (\sim (x = y))$': it doesn't matter whether we treat identity as logical or not when testing this sentence for logicality since whatever our choice of logical terms, this sentence is not a logical truth under representational semantics because there is a model making it false, namely a model with only one element.

What is needed, then, is an account of logical consequence which contains a modal ingredient. This modal ingredient is essential to handle the counter-examples studied in sections 4.10 and 4.11. The interpretational accounts, both Quine's and Tarski's, are attempts to define logical consequence without modalities. Presumably, this is what made these accounts attractive in the first place, but it is also why, finally, they are unsuccessful. It does not mean that a correct account has to be exclusively modal. Indeed, Shapiro (1998) and Hanson (1997) have both suggested that the best account of logical consequence consists in blend of the interpretational and the representational approaches.[30] Such a blended characterisation of logical consequence can be formulated thus:[31]

(BLD) Θ is a logical consequence of Γ iff Θ holds in all possibilities under every interpretation of the non-logical terminology in which Γ holds.

---

[30] Hanson calls the interpretational approach the 'formal approach' and the representational approach the 'modal approach'. We will use both terminology in what follows.

[31] This formulation is taken from Shapiro (1998: 148). Hanson's definition is virtually identical (1997: 379).

The preceding sections have been concerned mainly with making a case to the effect that Quine has no choice but to embrace a modal account such as (**BLD**). If this is right, then Quine cannot be an Error-theorist about logical necessity. He has to acknowledge the need, in order to explain and define the central concept of logical consequence, for a representational account making an irreducible use of modal notions. In the Appendix, we investigate the nature of the modality inherent in (**BLD**) in greater detail.

4.13   Conclusion

In this Chapter, as in the previous Chapters, the objective was to assess a brand of scepticism about logical necessity. We have seen, first, that Quine's account of logical truth must embrace a notion of meaning in order to specify the interpretation of the set $\mathfrak{S}$ of logical constants. However, Quine might be able to bypass that objection by putting forward an extensional account of meaning, even if it is not clear how that might be done. Sections 4.4–4.12, however, pointed to a much more serious defect in relation to his account of logical consequence. Indeed, we have seen that the substitutional account cannot be made to work — any theory of logical consequence has to be modal if it is to answer the counterexamples raised. Such an account does not have to be exclusively modal. (**BLD**), a blended account containing both an interpretational element and a modal element, does answer the counterexamples, and the modality in play allows us to represent different size of the domain. Even if the nature of the modality in (**BLD**) is minimal, it still shows its irreducibility, and this is enough for the purposes at hand. An interesting line of research for future work would be to assess whether the notion of modality in (**BLD**) is strong enough to account for all our intuitions about what follows from what — perhaps we need an even stronger notion of modality, one allowing us to represent not merely different sizes of the domain, but possibilities in general.

CHAPTER 5

CONCLUSION

5.1   Summary of Chapters 1–4

In Chapter One, we carefully analysed the argument in sections 1–4 of 'Two Dogmas'. This argument is often taken to be directed against the very intelligibility of notions such as analyticity, meaning, necessity and synonymy. If these were the argument's target, and if the argument had proved to be correct, that would have amounted to one of the most revolutionary philosophical argument ever produced. But is such an argument possible even in principle? Indeed, how can notions such as meaning and synonymy be unintelligible? If the very notion of an expression meaning something turned out to be spurious, that would mean, so to speak, that nothing makes sense. Such a conclusion itself does not make sense. Its very formulation requires expressions to have meaning. If Quine was arguing against the very notion of meaning, he was bound— prima facie at least — to fail. We have seen that the argument's target, properly interpreted, is not the notion of meaning qua notion, but only a particular understanding of meaning. Quine's charge against meaning is to be understood in the light of his aim to deconstruct the a priori via an attack on the analytic / synthetic distinction. This is to be explained by the fact that Carnap's notion of the a priori in *Logical Syntax* rests on the analytic / synthetic distinction. This distinction, in turn, rests on the notion of meaning since to be analytic is to be true by virtue of meaning (for Carnap). Viewed in this light, it becomes clear that Quine is only attacking the notion of meaning needed to form the building blocks of the analytic / synthetic distinction. The notion of meaning required to play this role, according to Quine, is the Platonistic, abstract notion of meaning. In short, Quine is arguing against the myth of the museum, in which meanings are the exhibits and words are labels for those abstract exhibits. This interpretation of Quine's argument does not imply scepticism about meaning in general, but, rather, only scepticism about the myth of the museum. I argued in Chapter One that Quine's argument, so construed, is compelling.

This study of Quine's argument against analyticity was important in the context of a thesis on scepticism about logical necessity. Indeed, for Carnap, what logically follows from what is to be decided on the basis of analytic statements. In Carnap's system, the notion of logical necessity is built upon the notion of analyticity. More precisely, it is because the rules of inference are analytic that we have the inexorability of statements of logical consequence. Accordingly, an argument against the intelligibility of analyticity as understood by Carnap amounts to an argument against the intelligibility of a notion of logical necessity based on analyticity.

The conclusion to be drawn from this study of what I have called 'Radical Irrealism' is twofold. Firstly, Quine's argument, when properly interpreted, is cogent and convincing. Secondly, the fact that Carnap's notion of logical necessity (according to Quine) requires the myth of the museum concerning meaning is good reason to make us suspicious of its viability. Correspondingly, we should also accept that a correct theory of logical necessity must avoid a commitment to Platonism about meaning. Such a commitment would weaken the plausibility of the theory, making it open to the forceful arguments advocated by Quine as well as other arguments against Platonism about meaning, such as those of Wittgenstein in his *Philosophical Investigations*.

In Chapter One, then, I am sympathetic to Quine's scepticism — so long as the argument is not inflated to amount to a full-blown attack on the very notion of meaning. In Chapter 2, I examined Quine's epistemological holism, according to which all statements have a relation to experience, even statements having to do with mathematics and logic. The difference between logical statements and, say, observational statements is one of degree, not kind. That is, no statements are situated outside the web of belief. There are no a priori statements, no analytic statements and no necessary statements. They all are a posteriori and contingent. However, Quine makes, in his philosophy of logic, a metatheorem / theorem distinction. This cleavage is not without its problems. By approaching the matter via the puzzle raised by Lewis Car-

roll in his 'What the Tortoise said to Achilles', we have seen that there is a tension in Quine's thoughts about the foundations of logic. He is, on the one hand, an advocate of naturalised epistemology, in which logic has no normative status; on the other hand, he is also a traditional logician, for whom logic encodes the norms of correct reasoning. I concluded that Quine's philosophy of logic does not have the resources to resolve this tension, leaving us with an unsatisfying account of the status of logic. Drawing on Davidson's views about weakness of the will, I then offered a solution to Carroll's puzzle contriving to be internalist in spirit while avoiding the excesses of externalism.

In Chapter 3, I examined a related argument against Quine's holism, one advocated by Wright. This argument purports to show that holism, to be a coherent epistemological picture, needs statements which are logically necessary. These will be precisely statements of the form $(\Gamma \vdash_L \Theta)$, that is, statements regulating the logical consequences in the web. After studying possible replies available to a Quinean, I contended that Wright's argument goes through and conclusively shows the need for logically necessary statements regulating the logical links in the web. Although Hale and McFetridge argue that there is a gap in Wright's argument, I suggest that once the target of Wright's argument is clearly identified and once we are clear about the nature of our judgements of logical necessity, there is a natural way to fill the gap. Thus, Wright's argument convincingly shows not only that Quine's philosophy of logic is unsatisfying, but also shows the necessity of logical necessity.

In Chapter 4, we examined another aspect of Quine's scepticism about logical necessity, namely his account of logical truth and logical consequence. Firstly, a case has been made to the effect that his theory must take into account the meaning of the logical constants, contrary to what he claims. This in turn shows that logical truths are true because they are analytic — true by virtue of meaning. However, this criticism does not directly address the issue of logical necessity since Quine can

cash out the required notion of meaning in a behaviourial framework. The issue then becomes whether such a notion is viable; but at least it forces Quine to acknowledge the need for a notion of meaning. Evaluating the viability of such a notion of meaning would have taken us too far from our main topic, so the question was left open. The remaining of the Chapter evaluated Quine's substitutional theory of logical consequence. After an investigation of the theory and the various problems it faces, I have concluded that Quine has to espouse a representational account of logical consequence, one in which modality plays an essential role (more will be said on the notion of modality build in the representational account in the Appendix). For the moment, let's point out that the result obtained holds generally, not only against Quine's substitutional account. Indeed, the counter-examples that have been raised against interpretational theories (including Tarski's) show the need for an irreducibly modal account. The crucial point is that a correct account of logical consequence has to allow for variation of domains. This is what, ultimately, forces any theory to include a modal element. In sum, the main thesis advanced in each of the Chapters is that scepticism about logical necessity is untenable. We need to acknowledge the inexorability of using logical necessity.

The Appendix can be seen as an example of future research on the topic of logical necessity. The strand explored has to do with the nature of modality: I will attempt to have a stronger grip on the modality in play in the blended account and also to assess its relation to various levels of discussion used when talking about inference. For now, however, I will indicate other avenues for future work.

5.2 From non-scepticism to a positive account

Now that a case has been made against scepticism about logical necessity, many questions remain. In fact, the implausibility of scepticism opens a realm of philosophical problems. We can now ask anew Dummett's questions: what is the source

of logical necessity, and how do we recognise it? In the Introduction, we saw that Dummett's questions presuppose that there is such a thing as logical necessity. Now that the matter has been resolved positively, it is worthwhile to notice that there is another presupposition in Dummett's questions, namely that logical necessity is something that we *recognise*. Accordingly, perhaps one of the most pressing problem is whether we should adopt a cognitivist or a non-cognitivist view about logical necessity. If the right account is the cognitivist one, then we can ask Dummett's questions. However, if we argue for a non-cognitivist account of logical necessity, the fundamental problem of logical necessity becomes: what exactly is the source of logical necessity? If it's not something we recognise, then the source has to come from us. How this metaphor should be cashed out is the main difficulty. Wittgenstein sometimes seems to think that what is logically necessary is a matter of decision. What does he mean exactly? Others advocate an expressivist view, according to which we 'project' our imaginative limitations onto the world. The projection metaphor raises a lot of problems — Blackburn, for one, is an expressivist who is cautious about its use (Blackburn, 1986: 122–123), preferring to talk about 'policy'. Hale (1989) and Wright (1980, 1989) also discuss the cognitivist / non-cognitivist debate, using the fictional figure of the Cautious Man to see whether there are facts of the matter about logical necessity.

Arguing for a cognitivist standpoint also raises a lot of difficulties. They are nicely spelled out by Dummett's questions: indeed, if we recognise logical necessity — if indeed it is something we 'track' — what exactly is the object of our cognition? Is it an abstract property? Do we recognise it by a special faculty of intuition? The ontological and epistemological puzzles raised by cognitivism are at least as considerable as those raised by non-cognitivism. Hale, in the aforementioned paper, criticises the viability of non-cognitivism and is tempted to adopt a cognitivist view on the matter — although, as he recognises, the epistemological and ontological

difficulties faced by such a view are considerable. While it is of course impossible to discuss these issues here, the cognitivism / non-cognitivism debate is perhaps the most fundamental one to be addressed with regards to logical necessity.

The discussion of Carroll's regress in Chapter 2 points to another set of problems, this time regarding statements of logical consequence and our inferential practice. Now that it has been established in the course of the thesis that true statements of logical consequence are necessary, we can enquire about our entitlement to draw a conclusion on the basis of the premises. How is our warrant to be understood? Does the fact that statements of logical consequence are necessary play a role in that entitlement? Should we adopt an externalist or an internalist stance about our inferential practice? In the section on Carroll's regress, I offered a solution congenial to an internalist outlook. The solution proposed could be understood in relation to Wittgenstein's rule-following considerations, especially the issues concerning the epistemology of rule-following and our understanding of rules, more precisely Wittgenstein's remarks that 'justification comes to an end' and that when one obeys a rule, one obeys it 'blindly'. My solution tries to steer a middle way between the excesses of externalism and the vulnerability of internalism to Carroll's regress. Accordingly, the heart of the solution has to be that when we infer by using modus ponens, we are following a norm — but this norm would have to be granted a special status, protecting it from the regress. I believe that Wittgenstein's remarks on rule-following contain the seeds of a convincing explanation congenial to my solution. This is, of course, another matter to be addressed in future work.

To sum up, three main lines of research are definitely worth pursuing in relation to the topics discussed in this dissertation. There is first the issues of the status and role of a modal account of logical consequence. This set of problems sprang from the discussion of Quine's and Tarski's account in Chapter 4. Secondly, the cognitivism / non-cognitivism debate about logical necessity is a natural extension of the conclusion

of this thesis. The last decade has seen few contributions on this debate — it deserves a revival. Maybe the relative silence is to be explained by the contemporary trend consisting of using modal discourse in the philosophical literature without enquiring about its nature. Indeed, nowadays, it seems that modal idioms are taken for granted. However, we should perhaps be unsatisfied about a philosophical theory taking for granted such a powerful tool — if the theory depends on something for which no account is available, then that theory's appeal is diminished. Finally, the third line of research is related to Carroll's regress. This puzzle forces us to be clear about the nature of our inferential practice. This hard question must be addressed: what is the nature of our warrant — if there is such a thing — when we correctly infer a conclusion from a set of premises?

APPENDIX A

APPENDIX

A.1 Quinean worries answered

It has been acknowledged that there is modality 'built into' our notion of logical consequence — this was indeed the main conclusion of the thesis. Now what remains to be done — amongst many other things — is to assess the nature of this modality. The problem can be put in those terms. We now know that in a valid argument, the conclusion necessarily follows from the premisses. That is, we have to 'box' statements of logical consequence. This is what logical necessity amounts to. At this stage of our enquiry, it becomes legitimate to ask about the meaning of this notion of logical necessity. That is, the purpose so far has been to stress the need of understanding true statements of the form $(\Gamma \vdash_L \Theta)$ as inevitably modal — but modal in which sense?

At this point one could legitimately raise the following worry. Quine does allow a role for logical necessity concerning validity. In a passage which appears puzzling, Quine writes (1953c: 169):

> There is no difficulty as long as necessity is construed as validity relative say to the logic of truth functions and quantification and perhaps classes.(...) As long as necessity in semantical applications is construed simply as explicit truth-functional validity (...) or quantificational validity, or set-theoretic validity, or validity of any well-determined kind, the logic of the semantical necessity predicate is a significant and very central strand of proof theory. But it is not modal logic, even unquantified modal logic.

Quine is here saying that necessity, understood as validity, is not problematic. Necessity thus understood is exactly what this thesis is about. Wouldn't it be possible, then, to reconcile the conclusions reached in the previous two Chapters with a Quinean understanding of logical necessity, so that the painstaking investigations leading to the need for a modal notion turned out to be deflated? To settle this question, we

have to reflect on the reasons why Quine takes himself to be entitled to say that necessity so construed is not problematic. His strategy involves two stages. Firstly, to reduce necessity to validity. Secondly, to explain validity with an interpretational account. For Quine, once we have a clear-cut account of validity, it is true that we have adequately explained logical necessity as it manifests itself in truth-functional logic or any logic for which we can define validity. This is because Quine will see no reason to reject a notion which satisfies his standards on definition and which can be shown to have its utility.

However, we have seen that the interpretational definition of validity — Quinean or Tarskian — cannot be made to work satisfactorily. An irreducibly modal element is needed. The exact nature of this modal element will be considered in the next sections, but nothing guarantees that this notion will lend itself to a definition satisfying Quine's high standards required for definitions. In sum, Quine relies on his account of logical consequence to justify the claim made in the quoted passage. But it was the point of Chapter 4 to show that this account does not work.

Another line of reply available to a Quinean is the following. Granted, she could say, that in the previous Chapters we have shown the need to incorporate modality in the concept of logical consequence. But the notion of necessity in question should be understood merely as a 'certainty' operator. That is, we ascribe necessity to correct statements of logical consequence because these statements figure at the centre of the web. They are considered 'true come what may' because they are deeply entrenched beliefs. As a result, we find it natural to accord them a special status and elevate them as holding of necessity. But, the argument goes, this is just a psychological fact about us, not reflecting anything in the way the world is. We could as well — this is a crucial bit — abandon those beliefs in logical necessity and it wouldn't make a difference. Once we realise the source of the belief in logical necessity — deeply entrenched belief — its superfluousness becomes apparent, and it should be seen as

a mere façon de parler, a rhetorical device.

Once again, this attempt to 'deconstruct' logical necessity fails. The Quinean is here claiming that belief in logical necessity reflects a confusion between beliefs that are situated at the centre of the web and beliefs that have a special status. The Quinean holds that no beliefs have a special status, they are all on the same level. It is because those beliefs are central to our conceptual scheme that we think, misleadingly, that they are logically necessary.[1] A proper analysis of those beliefs, the sceptical Quinean says, shows that we can do without them. But this line of argument does not work in light of the arguments of Chapter 3. Wright's argument shows that belief in logical necessity is not an option: it is in effect inexorable. Therefore, the Quinean move to reduce belief in logical necessity to mere deeply entrenched belief is blocked. We have to believe in the necessity of some statement of logical consequence: it is not an option or a display of a psychological attitude such as certainty. The challenge is to understand the modality in question. To try and reduce it to something less problematic simply won't work.

A.2   Talking about inference

Before exploring the question of the nature and the role of the modality in (**BLD**), we need to clear up the field. In particular, there is a need to discern between different levels of discussion used when talking about inference. What, for instance, is the relation between a characterisation of logical consequence such as (**BLD**) and first-order logic? How does first-order logic relate to our intuitive notion of 'what follows from what'? How does the latter relate to (**BLD**)? What is first-order logic intended

---

[1]Those beliefs might have a dialectical, pragmatical role. This sense is manifested when we want to express an almost unshakeable belief in something. For instance, someone might say that the conclusion 'necessarily' follows from a set of premises in order to emphasise the apparent inexorability, but she need not believe in necessity in a stronger sense than that. This 'conversational' sense of necessity is not the one which is our concern here.

to model? We are in a conceptual muddle. In what follows, I will distinguish between three levels of discussion we use when talking about inference.

The first level is the *intuitive* one. This is a very primitive level of discussion. When talking about inference, we first have in mind the notion of 'what follows from what'. This is the most general characterisation of the topic under study in this thesis. But as such, it is not very illuminating: it refers, roughly, to our practice of inferring conclusions from premises. In this sense, this example illustrates a correct instantiation of the concept of 'what follows from what':

(1) All men are mortal

(2) Socrates is a man

(3) Therefore, Socrates is mortal.

No one can really dispute that. But the following example also instantiates this intuitive characterisation:

(4) Al is taller than Bill

(5) Therefore, Bill is shorter than Al.

At the intuitive level, (5) follows from (4). Indeed, in actual reasoning, it is correct to infer (5) from (4), unless one is doing things at a theoretical level — this argument is not valid in first-order logic, since there are arguments displaying the same form in which the premise is true and the conclusion false. But in normal discourse, no one would object in the following way: "Actually, you cannot infer that Bill is shorter than Al from the premise that Al is taller than Bill. I do not accept this inference, and if you want to persuade me that (5) follows from (4), you will need to do better than that." Such a person would appear somewhat eccentric.

Staying at the intuitive level, however, one could grant that (5) follows from (4), but that it does not *logically* follow. Such a person would say that logical consequence is a matter of form, not of content. As such, the second example is not valid in virtue of form, whereas the first is. Here we witness a restriction in the notion of 'what follows from what' at the intuitive level. Both examples above are correct instances of what follows from what, but only the first displays a correct logical consequence.

Next comes the *pre-theoretical* level of discussion. At that level, we try to articulate the intuitive notions in more precise terms. For instance, a straightforward pre-theoretical characterisation of the unrestricted notion of 'what follows from what' (the one embracing both examples above) would look like this:

(MOD) Θ follows from Γ if and only if it is impossible for Γ to be true and Θ to be false.

It is clear that such a characterisation captures the intuitive notion of what follows from what. Importantly, this is essentially a modal representation of the intuitive notion. Turning now to the restricted notion — validity in virtue of form — we have seen in Chapter 4 that there are different ways to characterise it in pre-theoretic terms. First we studied the substitutional characterisation:

(SUB) Θ is a logical consequence of Γ if and only if there is no substitution of the non logical terms in which Γ is true but Θ is false.

There was also Tarski's account:

(TLC) Θ is a logical consequence of Γ if and only if every model of the sentences in Γ is also a model of Θ.

Tarski's account, as we have seen, can be considered as an attempt to capture the restricted notion only if the notion of model is understood in the interpretational way — otherwise it's a version of (MOD). Finally, at the end of last Chapter, we have

seen that a blended account might be the best we have to model validity in virtue of form:

(BLD) Θ is a logical consequence of Γ if and only if Θ holds in all possibilities under every interpretation of the non-logical terminology in which Γ holds.

Before examining the various relations between these characterisations and what they are intended to capture, there is also a third level of discussion to consider, namely the *theoretical* level. At the theoretical level we have the logical systems, whose role is to model, in a technical sense, our reasoning practice. For instance, classical first-order logic is meant to capture the intuitive idea that validity is a matter of form. In contrast, we could say that relevant logics and modal logics are concerned to model the unrestricted notion of what follows from what — these logics use modal concepts that are meant to model the modality in **(MOD)**.

To recapitulate, we have in this section identified three levels of discussion we use when talking about inference. The first level refers to our rather vague intuitions about inference and what follows from what. In the second level — the pre-theoretical one — these intuitions are articulated more precisely. The third level is the theoretical one. In this level we find the logical systems. We can identify two main strands going through the levels. There is first the strand intended to capture the notion of *logical* consequence, that is, the notion of validity in virtue of form. At the first level, it corresponds to the restricted notion of what follows what. Second level we find its associated pre-theoretical notions, such as **(SUB)**, **(TLC)**, **(BLD)**. Corresponding to these pre-theoretical notions we have, for instance, first-order logic, which belongs to the third level. The other main strand starts with the unrestricted notion of what follows from what. A pre-theoretic account of this notion would look like something along the lines of **(MOD)**. At the theoretical level, we find various modal logics attempting to represent **(MOD)**.

In the next sections, we will use this tripartite framework to study the relations

between these different characterisations. The pressing question, for present purposes, is to enquire into whether the minimal sense of modality in (**BLD**) is the best we can do — whether, that is, it can be argued that a full-blown notion of modality such as the one displayed in (**MOD**) is required to account for validity.

A.3 Kreisel's proof and (BLD)

The objective in the remaining of the Appendix is to assess the nature of the modality in (**BLD**). What has been established, so far, is that there is modality in the notion of logical consequence at the pre-theoretical level. Now we can ask: what is the relation between (**BLD**) and, on the one hand, our intuitive ideas about 'what follows from what', and, on the other hand, logical systems? As for the first question, the answer is straightforward: (**BLD**) is designed to capture our formal intuitions about logical consequence, not the unrestricted notion of what follows from what. The modality in (**BLD**) is there to handle counterexamples having to do with the size of the domain. This point is important — the notion of possibility in (**BLD**) has to be understood strictly as representing different sizes of the domain, nothing else. This is why it is a 'modest' notion of modality and not a 'full-blown' one like the one that appears in (**MOD**). If the modality in (**BLD**) were playing the role of representing possibilities in an unrestricted sense, the account would be undistinguishable from (**MOD**) — it wouldn't be an account of logical form anymore. In sum, (**BLD**) is intended to capture the insight that logical consequence is a matter of form and its modality has to be understood in a restricted sense.

Now what about the relation between (**BLD**) and the theoretical level? Is there a logic capable of representing inferential practice so as to capture the insight that logic is a matter of form and to deal with its irreducible notion of modality? Indeed, we must not forget that (**BLD**) is only a general formulation of what a correct, formalised logical system must capture. In particular, we want to know whether

standard first-order model theory (let's call it (**SMT**)) is a good model, so to speak, of (**BLD**). There is currently much work being done in that field.[2] Interestingly, perhaps the most ground-breaking results in this direction were obtained more than three decades ago by Georg Kreisel in a few paragraphs of his paper 'Informal Rigour and Completeness Proofs'. To answer the question at hand, I suggest the following strategy: that we use the well-known argument from Kreisel showing that (**SMT**) captures our intuitive notion of validity in virtue of form. This in turn will be sufficient to 'prove' that (**SMT**) succeeds in modelling (**BLD**). I turn now to Kreisel's proof.

Following Kreisel, let '$Val\alpha$' means that α is intuitively valid — valid in all structures — and '$V\alpha$' means that α is model-theoretically valid (*i.e.*, valid in the sense of (**SMT**)) — valid in the set-theoretic structure. Notice that 'α' can be either a single sentence (a logical truth) or can consist in a set of sentences (a logical consequence). Kreisel's proof shows that V and Val are co-extensional, that is, that:

$$\forall\alpha(Val\alpha \equiv V\alpha) \tag{A.1}$$

is true for first-order logic. The left to right conditional:

$$\forall\alpha(Val\alpha \rightarrow V\alpha) \tag{A.2}$$

is intuitively true. To see why, let's contrapose it:

$$\forall\alpha(\sim V\alpha \rightarrow \sim Val\alpha) \tag{A.3}$$

This is true for the following reason. Let's imagine that an argument is set-theoretically invalid. That means that there is a set-theoretic model in which the premisses of this

---

[2]See Shapiro (1998: 147-155), Hanson (1997: 398-407), and Sher (1991) for more on the relation between model-theoretic logical consequence and the intuitive notion of logical consequence.

argument are true and the conclusion false. That in turns means that the argument is invalid in *a* structure, thus showing that it can't be intuitively valid since to be intuitively valid means to be valid in *all* structures. The right to left conditional:

$$\forall \alpha (V\alpha \rightarrow Val\alpha) \tag{A.4}$$

requires a bit more work. The proof uses Gödel's completeness theorem. Let '$D\alpha$' means 'α is formally derivable by means of some accepted set of formal rules' (see Kreisel (1967: 89)). By Gödel's completeness theorem (remember we're considering these relations only in the context of first-order logic), we have:

$$\forall \alpha (V\alpha \rightarrow D\alpha) \tag{A.5}$$

Now Kreisel offers an informal argument to the effect that:

$$\forall \alpha (D\alpha \rightarrow Val\alpha) \tag{A.6}$$

thus completing the equivalence proof between *V* and *Val*. The argument in support of '$\forall \alpha (D\alpha \rightarrow Val\alpha)$' is that the standard formal rules of inference in first-order logic (modus ponens, substitution rules, etc) are almost universally recognised as being valid. Kreisel: "...one reasons in mathematical practice, using the notion of consequence or of logical consequence, freely and surely" (1967: 90). In fact, in an axiomatic presentation of first-order logic, it is possible to use only two rules of inference, namely modus ponens and the rule of substitution. These two rules are valid if anything is. So since these two rules are, in theory, sufficient to formalise first-order logic, and since they are intuitively valid, we have strong grounds to hold true the theorem '$\forall \alpha (D\alpha \rightarrow Val\alpha)$'.

The proof shows the extensional equivalence between a notion at the intuitive level

and a logical system, which belongs to the theoretical level. As such, it does not say anything about (**BLD**) — belonging to the pre-theoretical level — and its relation to (**SMT**). However, (**BLD**) is designed explicitly to capture the intuitive idea that validity is a matter of form. Other pre-theoretical accounts, for instance (**SUB**), are also designed to capture the same intuitive notion, but the notion of substitution does not work : it leads to accept as valid arguments which are not valid. To handle these counterexamples, a pre-theoretical account has to allow for variations of domain, and this is the role of the modality in (**BLD**). The idea behind (**BLD**) is, I take it, the following: we want to say that validity is a matter of form *and* independent of the size of the universe. The second clause, despite the fact that it leads one to incorporate a modal element in one's pre-theoretical account, is formal in spirit. Indeed, to claim that validity should not depend on the size of the domain amounts to claiming that validity has nothing to do with the way the world is — it shouldn't matter, if validity is strictly formal, how many objects there are. Admittedly, there is an irony here: to save the insight that logical consequence is strictly a matter of form, we have to include a modal element. In any case, (**BLD**) corresponds strictly to what we mean by validity in virtue of form. Since the latter notion is extensionally equivalent to (**SMT**), we have by transitivity that (**BLD**) and (**SMT**) are also extensionally equivalent.

This result is welcome, since it means that the current model-theoretic definition already in place is a good model of what we mean by logical consequence. A Quinean, I contended in Chapter 4, is forced to embrace something like it for first-order logic. For second-order logic, no one has yet been able to prove the same result. It is bound to be more complicated, since for higher-order logics the completeness theorem does not hold and the proof for the first-order case makes essential use of it. But a strong Quinean will not be bothered by that, given that for her second-order logic is not logic.

However, there is a sense in which it might be welcomed by a Quinean. In effect, Kreisel's proof demonstrates an equivalence between two notions not containing any modal element. The proof bypasses the pre-theoretical level, and this is the level where modality appears. Indeed, in our intuitive notion of logical form, it seems that there is no modality involved. Prima facie, 'truth in all structures' contains no mention of modality. Same for (SMT): it is certainly not modal logic. What happened to the modality involved in the blended account? It's as if (SMT) had sublimated the modality, managing to be both a representation of (BLD) and free of modal notions. We can rest assured that (SMT) will represent the modality in the blended account assuming that the models provide domains of different size.

Nonetheless, it is not that clear that a Quinean can sensibly applaud the fact that (SMT) is not a modal logic. What (SMT) does is to model our inferring practice — more precisely, it models our notion of validity in virtue of form. A pre-theoretical account, such as (BLD) or (SUB), is meant to clarify the intuitive notion of validity, that is, to make sense of it in an illuminating and explanatory way. Since it has been shown in the last Chapter that any pre-theoretical account of logical form has to contain a modal element, we can safely say that there is an irreducible modal element in the concept of logical consequence. That (SMT) satisfyingly models the blended account without using modality is only because there are enough sets to model any size of domain. Put differently, one could say that (SMT) has modality built right into its notion of models — they are playing the representational role. Admittedly, the possible models represent only various size of the domain — but it does, in the end, model modality. If there is a relief to be found in the Quinean camp, it should rather be that the modality in (BLD) is not full-blown.

A.4 Modal intuitions about inference

(**BLD**) captures our formal intuitions about what follows from what. It does not capture the unrestricted notion of what follows what, the one according to which we can infer that Bill is smaller than Al if it is true that Al is taller than Bill. We want to say that the conclusion does not logically follow from the premise in that kind of case because there are arguments with the same logical form which have a true premise and a false conclusion. Indeed, 'taller than' and 'smaller than' are not logical terms. So it is easy to replace those non-logical terms with other terms rendering the argument false. But as a matter of practice, it is correct to infer that Bill is smaller than Al from the given premise. We infer in the unrestricted sense all the time — it works well, it's an established practice. Accordingly, what has to be conceded is that (**BLD**) and(**SMT**) do not accurately model the way we actually infer. They undergenerate in that sense.

This seems unfortunate. The splitting of the notion of what follows from what between its restricted sense and its unrestricted sense appears completely artificial from the standpoint of our actual practice of inferring. That does not mean, of course, that the restricted notion is not interesting, nor that (**SMT**) is not highly important and fruitful. However, it does remain the case that (**SMT**) is limited as a model of inference. Indeed, let us imagine someone who does not know how to infer, or who, for some reason, infers very erratically. We then successfully teach that person (**SMT**). Then we confront that person with the premise 'Al is taller than Bill' and we ask her if she would infer from that the sentence 'Bill is shorter than Al'. A good logic student, she argues that she can't draw that inference, because she formalised that argument into the language of (**SMT**) and it showed that it has an invalid argument-form. Thus, for any intuitively correct inference that is not valid in virtue of form, she rejects it — not only in the classroom or in her logic tests, but also in her day-to-day life. Since so much of our actual reasoning is not strictly formal, such

a person would still struggle outside the classroom despite having learned everything about (**SMT**). This shows that (**SMT**) is a limited tool to model reasoning.

There is no doubt that on the pre-theoretical level, a purely modal account such as (**MOD**) captures the unrestricted notion. There is also no doubt that on the theoretical level modal logics are intended to model (**MOD**) and the unrestricted notion. It is tempting to conclude that any correct logic whose role is to represent the way we actually infer would have to be modal. To a Quinean who would agree that logic's main role is to represent the way we actually infer, this conclusion would be totally catastrophic, yet unavoidable — a commitment to both (**MOD**) and modal logic is a Quinean's nightmare but an anti-sceptic's dream. This, if correct, represents an important victory for the defender of modalities.

REFERENCES

[Barrett and Gibson (1990)] Barrett, R., and R. Gibson, eds. (1990) *Perspectives on Quine*, Blackwell.

[Blackburn (1986)] Blackburn, S. (1986) "Morals and Modals," in G. Macdonald and C. Wright, eds., *Fact, Science, & Morality*, Blackwell, 119–141.

[Blackburn (1995)] Blackburn, S. (1995) "Practical Tortoise Raising," *Mind* 104, 695–711.

[Boghossian (1997)] Boghossian, P. (1997) "Analyticity," in B. Hale and C. Wright, eds., *A Companion to the Philosophy of Language*, Blackwell, 331–368.

[Bolzano (1837)] Bolzano, B. (1837) *Wissenschaftslehre*, Meiner. Translated into English by R. George in 1972.

[Boolos (1975)] Boolos, G. (1975) "On Second-Order Logic," in *Logic, Logic, and Logic*, Harvard University Press, 37–53.

[Boolos (1984)] Boolos, G. (1984) "To Be is to Be a Value of a Variable," in *Logic, Logic, and Logic*, Harvard University Press, 54–72.

[Boolos (1998)] Boolos, G. (1998) *Logic, Logic, and Logic*, Harvard University Press.

[Butterfield (1986)] Butterfield, J., ed. (1986) *Language, Mind, and Logic*, Cambridge University Press.

[Carnap (1937)] Carnap, R. (1937) *The Logical Syntax of Language*, Routledge & Kegan Paul.

[Carnap (1956)] Carnap, R. (1956) *Meaning and Necessity*, Chicago University Press, second edition.

[Carnap (1957)] Carnap, R. (1957) "Empiricism, Semantics, and Ontology," in *Meaning and Necessity*, Chicago University Press, second edition.

[Carroll (1895)] Carroll, L. (1895) "What the Tortoise said to Achilles," *Mind* 4, 278–280.

[Coffa (1991)] Coffa, A. (1991) *The Semantic Tradition from Kant to Carnap*, Cambridge University Press.

[Davidson (1970)] Davidson, D. (1970) "How is Weakness of the Will Possible?" in *Essays on Actions and Events*, Clarendon Press, 21–42.

[Davidson (1982)] Davidson, D. (1982) "Paradoxes of Irrationality," in J. Hopkins and R. Wollheim, eds., *Philosophical Essays on Freud*, Cambridge University Press, 289–305.

[Dummett (1959)] Dummett, M. (1959) "Wittgenstein's Philosophy of Mathematics," in *Truth & Other Enigmas*, 166–185.

[Dummett (1978)] Dummett, M. (1978) *Truth & Other Enigmas*, Harvard University Press.

[Dummett (1991)] Dummett, M. (1991) *The Logical Basis of Metaphysics*, Harvard University Press.

[Etchemendy (1990)] Etchemendy, J. (1990) *The Concept of Logical Consequence*, Harvard University Press.

[Frege (1884)] Frege, G. (1884) *The Foundations of Arithmetic*, Northwestern University Press. Translated by J. Austin.

[Grice and Strawson (1956)] Grice, P. H., and P. F. Strawson (1956) "In Defence of a Dogma," *Philosophical Review* 65, 141–158.

[Hale (1989)] Hale, B. (1989) "Necessity, Caution,and Scepticism," *Proceeedings of the Aristotelian Society* 63, 175–202.

[Hale (1996)] Hale, B. (1996) "Absolute Necessities," in *Philosophical Perspectives*, volume 10, Blackwell, 93–117.

[Hale (1999)] Hale, B. (1999) "On Some Arguments for the Necessity of Necessity," *Mind* 108, 23–52.

[Hanson (1997)] Hanson, W. (1997) "The Concept of Logical Consequence," *Philosophical Review* 106, 365–409.

[Hookway (1988)] Hookway, C. (1988) *Quine: Language, Experience and Reality*, Polity Press, Blackwell.

[Jackson (1998)] Jackson, F. (1998) *From Metaphysics to Ethics: A Defence of Conceptual Analysis*, Blackwell.

[Kreisel (1967)] Kreisel, G. (1967) "Informal Rigour and Completeness Proofs," in J. Hintikka, ed., *The Philosophy of Mathematics*, Oxford University Press, 78–94.

[Kripke (1971)] Kripke, S. (1971) "Identity and Necessity," in M. Munitz, ed., *Identity and Individuation*, New-York University Press, 135–164.

[Kripke (1980)] Kripke, S. (1980) *Naming and Necessity*, Harvard University Press.

[Kripke (1982)] Kripke, S. (1982) *Wittgenstein on Rules and Private Language*, Blackwell.

[Lazar (1999)] Lazar, A. (1999) "Akrasia and the Principle of Continence or What the Tortoise Would Say to Achilles," in L. E. Hahn, ed., *The Philosophy of Donald Davidson*, Open Court, 381–401.

[McFetridge (1990)] McFetridge, I. (1990) "Logical Necessity: Some Issues," in J. Haldane and R. Scruton, eds., *Logical Necessity & Other Essays*, Aristotelian Series, 135–154.

[Putnam (1983)] Putnam, H. (1983) "Possibility and Necessity," in *Philosophical Papers*, volume 3, Cambridge University Press, 46–68.

[Quine (1936a)] Quine, W. V. O. (1936a) "Set-theoretic Foundations for Logic," in *Selected Logic Papers*, 83–99.

[Quine (1936b)] Quine, W. V. O. (1936b) "Truth by Convention," in *The Ways of Paradox*, 70–99.

[Quine (1940)] Quine, W. V. O. (1940) *Mathematical Logic*, Harvard University Press, third edition.

[Quine (1941)] Quine, W. V. O. (1941) "Whitehead and the Rise of Modern Logic," in *Selected Logic Papers*, 3–36.

[Quine (1948)] Quine, W. V. O. (1948) "On What there Is," in *From a Logical Point of View*, 1–19.

[Quine (1950)] Quine, W. V. O. (1950) *Methods of Logic*, Henry Holt & Company, revised edition.

[Quine (1951)] Quine, W. V. O. (1951) "Two Dogmas of Empiricism," in *From a Logical Point of View*, 20–46.

[Quine (1953a)] Quine, W. V. O. (1953a) *From a Logical Point of View*, Harvard University Press.

[Quine (1953b)] Quine, W. V. O. (1953b) "The Problem of Meaning in Linguistics," in *From a Logical Point of View*, Harvard University Press, 47–64.

[Quine (1953c)] Quine, W. V. O. (1953c) "Three Grades of Modal Involvement," in *The Ways of Paradox*, 156–174.

[Quine (1960)] Quine, W. V. O. (1960) *Word & Object*, Mit Press.

[Quine (1963)] Quine, W. V. O. (1963) "Carnap and Logical Truth," in *The Ways of Paradox*, 100–125.

[Quine (1966a)] Quine, W. V. O. (1966a) *Elementary Logic*, Harvard University Press, revised edition.

[Quine (1966b)] Quine, W. V. O. (1966b) *The Ways of Paradox*, Random House.

[Quine (1969a)] Quine, W. V. O. (1969a) "Epistemology Naturalized," in *Ontological Relativity and Other Essays*, Columbia University Press, 69–90.

[Quine (1969b)] Quine, W. V. O. (1969b) *Set Theory and its Logic*, Harvard University Press, revised edition.

[Quine (1981)] Quine, W. V. O. (1981) *Theories and Things*, Harvard University Press.

[Quine (1986)] Quine, W. V. O. (1986) *Philosophy of Logic*, Harvard University Press, second edition.

[Quine (1992)] Quine, W. V. O. (1992) *Pursuit of Truth*, Harvard University Press.

[Quine (1995a)] Quine, W. V. O. (1995a) *From Stimulus to Science*, Harvard University Press.

[Quine (1995b)] Quine, W. V. O. (1995b) *Selected Logic Papers*, Harvard University Press, second enlarged edition.

[Quinton (1990)] Quinton, A. (1990) "Doing without Meaning," in R. Barrett and R. Gibson, eds., *Perspectives on Quine*, Blackwell, 294–308.

[Read (1995)] Read, S. (1995) *Thinking about Logic*, Oxford University Press.

[Schirn (1998)] Schirn, M., ed. (1998) *The Philosophy of Mathematics Today*, Clarendon Press.

[Sellars (1956)] Sellars, W. (1956) *Empiricism and the Philosophy of Mind*, Harvard University Press.

[Shapiro (1991)] Shapiro, S. (1991) *Foundations without Foundationalism: A Case for Second-Order Logic*, Oxford University Press.

[Shapiro (1998)] Shapiro, S. (1998) "Logical Consequence: Models and Modality," in M. Schirn, ed., *The Philosophy of Mathematics Today*, Clarendon Press, 131–156.

[Shapiro (2002)] Shapiro, S. (2002) "The Status of Logic." Forthcoming.

[Sher (1991)] Sher, G. (1991) *The Bounds of Logic*, MIT Press.

[Sher (1996)] Sher, G. (1996) "Did Tarski commit 'Tarski's fallacy'?" *Journal of Symbolic Logic* 61, 653–686.

[Strawson (1957)] Strawson, P. (1957) "Propositions, Concepts, and Logical Truth," *Philosophical Quarterly* 7, 15–25.

[Strawson (1959)] Strawson, P. (1959) *Individuals*, Methuen.

[Strawson (1985)] Strawson, P. (1985) *Skepticism and Naturalism: Some Varieties*, Methuen.

[Stroud (1984)] Stroud, B. (1984) *The Significance of Philosophical Scepticism*, Oxford University Press.

[Tarski (1933a)] Tarski, A. (1933a) "The Concept of Truth in Formalized Languages," in *Logic, Semantics, Metamathematics*, 152–278.

[Tarski (1933b)] Tarski, A. (1933b) "Some Observations on the Concepts of ω-Consistency and ω-Completeness," in *Logic, Semantics, Metamathematics*, 279–295.

[Tarski (1935)] Tarski, A. (1935) "On the Concept of Logical Consequence," in *Logic, Semantics, Metamathematics*, 409–420.

[Tarski (1983)] Tarski, A. (1983) *Logic, Semantics, Metamathematics*, Hackett, second edition. Edited and introduced by John Corcoran.

[Tarski (1986)] Tarski, A. (1986) "What are Logical Notions?" *History and Philosophy of Logic* 7, 143–154.

[Tennant (1997)] Tennant, N. (1997) *The Taming of the True*, Oxford University Press.

[Wittgenstein (1953)] Wittgenstein, L. (1953) *Philosophical Investigations*, Blackwell.

[Wittgenstein (1978)] Wittgenstein, L. (1978) *Remarks on the Foundations of Mathematics*, MIT Press, revised edition.

[Wright (1980)] Wright, C. (1980) *Wittgenstein on the Foundations of Mathematics*, Harvard University Press.

[Wright (1983)] Wright, C. (1983) *Frege's Conception of Numbers as Objects*, Aberdeen University Press.

[Wright (1986)] Wright, C. (1986) "Inventing Logical Necessity," in J. Butterfield, ed., *Language, Mind, and Logic*, Cambridge University Press, 187–209.

[Wright (1989)] Wright, C. (1989) "Necessity, Caution and Scepticism," *Proceedings of the Aristotelian Society* 63, 203–209.

[Wright (1992)] Wright, C. (1992) *Truth & Objectivity*, Harvard University Press.

[Wright (1997)] Wright, C. (1997) "On the Philosophical Significance of Frege's Theorem," in R. Heck, ed., *Language, Thought, and Logic*, Clarendon, 201–244.

[Wright (2002)] Wright, C. (2002) "On Basic Logical Knowledge: Reflections on Paul Boghossian's "How Are Objective Epistemic Reasons Possible?"." Forthcoming in *Philosophical Studies*.