

## INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

**The quality of this reproduction is dependent upon the quality of the copy submitted.** Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

**UMI<sup>®</sup>**

Bell & Howell Information and Learning  
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA  
800-521-0600



A

DEFLATING DEFLATIONISM

by

BRADLEY ARMOUR-GARB

A dissertation submitted to the Graduate Faculty of Philosophy in partial fulfillment of the requirements for the degree of Doctor of Philosophy, The City University of New York

1999

**UMI Number: 9946134**

**Copyright 1999 by  
Armour-Garb, Bradley Philip**

**All rights reserved.**

---

**UMI Microform 9946134  
Copyright 1999, by UMI Company. All rights reserved.**

**This microform edition is protected against unauthorized  
copying under Title 17, United States Code.**

---

**UMI**  
**300 North Zeeb Road**  
**Ann Arbor, MI 48103**

© 1999

BRADLEY ARMOUR-GARB

All Rights Reserved

## Approval Page

This manuscript has been read and accepted for the Graduate Faculty in Philosophy in satisfaction of the dissertation requirement for the degree of Doctor of Philosophy.

9/2/99

Date

9/2/99

Date



Chair of Examining Committee



Executive Officer

Stephen Schiffer

Arnold Koslow

Jonathan Adler

Simon Blackburn

---

Supervisory Committee

THE CITY UNIVERSITY OF NEW YORK

## Abstract

## DEFLATING DEFLATIONISM

by

Bradley Armour-Garb

Advisor: Jerrold J. Katz

In this dissertation, I take a close look at the deflationary theory of truth, and deflationary semantics, generally. My thesis is that, as a theory about the nature and function of the property of truth, deflationism is well supported. However, deflationary semantics, which combines deflationism about truth with deflationism about meaning cannot be argued for by pointing to the expressive function of the truth predicate.

Having shown that deflationism about meaning cannot be argued for in this way, I develop a challenge to deflationary semantics, the challenge of *the contingency of sentential truth conditions*. The challenge for the deflationist is to explain (or explain away) the fact that her view assigns the wrong modal properties to “true”. As I show, the deflationist does not have the resources either to explain or to explain away this counter-intuitive feature of her view. Since, in addition, the arguments for deflationism do not support the adoption of a deflationary attitude towards meaning, I offer and support what I call *propositional deflationism*. Propositional deflationism deflates truth for propositions, thereby respecting the deflationist’s contentions about the nature and function of truth, without deflating truth for sentences and beliefs. The result is a hybrid view that is deflationist about truth, though not about meaning.

## Acknowledgments

I am both attracted to and repelled by the deflationary conception of truth, a combination of attitudes that is reflected in the position that I develop in this dissertation. These conflicting attitudes can be traced to those of the two philosophers who, more than anyone else, have shaped the way I do philosophy—Jerry Katz and Stephen Schiffer. I would like to thank them for their time, patience and encouragement, and for teaching me how to become a philosopher.

I cannot adequately express the gratitude I feel towards these two men. Stephen was the reason that I came to the Graduate Center, and, to some extent, the reason that I entered philosophy in the first place. He taught me—largely by example—how to approach philosophical issues and philosophical positions; I do philosophy in the way that I do it largely because of his influence. I would like to thank him for his time, his patience, and his insight. I would not be here, were it not for Schiffer.

I came to work with Jerry a little later in my graduate school career, and it was easily the best decision that I made as a graduate student. Jerry's comprehensive picture of philosophy and his intellectual honesty and integrity have taught me what it is to be a philosopher. I would like to thank Jerry for the innumerable lunches and coffees and, more important, for inviting me into his philosophical circle and allowing me to work through problems—mine and his—with him. I am very lucky to have been able to work so closely with him, and I expect to continue to do so in the years to come.

I was also very lucky to get to work closely with two other very gifted philosophers, Jonathan Adler and Arnie Koslow. Although different in method and focus, Jonathan and Arnie are two of the most careful philosophers that I have ever encountered, and I have learned a tremendous amount from each of them. I would like to thank Jonathan for the great conversations we have had about truth and epistemology. More affecting, however, was his willingness to treat me as a colleague. Jonathan has helped me to develop philosophical confidence, and I will forever be indebted to him for that.

I would like to thank Arnie for indulging me in innumerable conversations about truth and logic. Early in my graduate school career, and up through the writing of my dissertation, Arnie taught me that doing logic was doing philosophy. He has taught me the importance of employing formal methods in my work and, more than he realizes, has helped to chart the path of my future research. Arnie's enthusiasm for philosophy and philosophical discussion really helped to make graduate school the incredible experience that it was, and helped to foster in me a love for the discipline.

I would also like to thank Jared Blank, David Pitt, Mark Sheehan, and James Woodbridge. Jared read and commented on a draft of the dissertation, and he has given me lots of great criticism over the last few years. David, who does not understand my fascination with deflationism, has been a fantastic philosophical friend and colleague, even if we rarely agree on anything. Mark has kept me sane during the writing of this dissertation, and (together with his wife Debbie) has become one of my best friends in the world. Whenever I was stuck on a problem or fed up with a section, I could call Mark, discuss the issues, and return to work, feeling fired up and excited to continue.

Although I became friends with James near the end of my dissertation, he, more than anyone else, has helped me to understand the intricacies of deflationism. I have benefited a tremendous amount from our philosophical discussions and from our friendship, and look forward to a future of talking philosophy.

Finally, I would like to thank my family: my parents, Andy and Sheila, my sister, Ali, my brother, Jonah, and my wife, Allison. Although I am frequently teased about philosophy (especially by Jonah), I have received so much support and love from them that I would like to dedicate this dissertation to them. No words I can think of can express my gratitude.

More than anyone else, I would like to thank my wife, Allison. She has been the most important person in my life and I do not know what I, or this dissertation, would be like without her. Allison's encouragement, trust and commitment have helped to make me, and my work, as good as can be, and, for that, and for making my life so wonderful, I thank her.

## Table of Contents

Chapter 0: Introduction .....	1
0.1 Deflationary Truth and Semantic Deflationism .....	3
0.2 Confusion Over the Deflationist's Claim .....	6
0.3 My Thesis .....	14
Chapter 1: The History of Deflationism .....	18
1.1 A Brief History of Deflationism .....	19
1.2 The Redundancy Theory.....	23
1.2.1 Ayer's Redundancy Theory.....	27
1.2.2 Strawson's Redundancy Account.....	29
1.2.3 Grover, Camp and Belnap.....	30
1.3 Frege's Anti-Correspondence Argument.....	32
1.4 The Quinian Theory.....	36
1.5 A Tarskian Deflationary Theory.....	43
1.6 Comparison of a Tarskian and a Disquotational View .....	56
1.7 Concluding Remarks .....	59
Chapter 2: Contemporary Deflationism.....	60
2.1 Regarding (T1) .....	63
2.1.1 Deflationary Theories vs. Tarskian Theories.....	64
2.1.2 The Status of the Instances of (DS) and Semanticalism .....	67
2.2 Regarding (T2) & (T3) .....	73
2.3 Regarding (T4) .....	77
2.4 Disquotationalism vs. Deflationism Revisited .....	79
2.5 A Problem with Idiolect Disquotationalism .....	83
2.6 Field, Substitutional Quantification and Schemata .....	96
2.7 More on Propositional Deflationism .....	101
2.8 What is it to Explain Truth? (One view) .....	103
2.9 Semanticalism .....	107
2.9.1 Idiolect Disquotationalism and Semanticalism Revisited .....	109
2.9.2 Property Deflationism and Semanticalism.....	113
Chapter 3: Explanatory Liaisons.....	117
3.1 Explaining Truth and the Explanatory Role of Truth.....	118
3.2 Explanatory Role? .....	120
3.3 Truth, Explanation and Deflationism.....	123
3.4 What is an Explanatory Liaison?.....	129
3.4.1 Explanatory Liaisons of Type I: Explaining Behavior.....	131
3.4.2 Introduction to Gibbons' Argument.....	132
3.4.3 Gibbons' Argument: Truth in Action.....	133

3.4.4 The Deflationist Response to Gibbons' Argument .....	139
3.4.5 A General Moral Regarding Psychological Explanation and Truth .....	146
Chapter 4: Truth and Meaning .....	149
4.1 Why Meaning Matters to the Deflationist: Jackson, Oppy and Smith .....	156
4.2 Horwich's Account, Briefly Summarized .....	162
4.2.1 The General Issue and the General Problems .....	167
4.2.2 Our First General Problem with Horwich's Account of Meaning .....	171
4.2.3 Our Second General Problem: Horwich and the Problem of Intentionality ..	177
4.3 Deflationist Meaning .....	187
4.3.1 A Deflationist Theory of Meaning? .....	187
4.3.2 A Lacuna in the View .....	190
Chapter 5: The Contingency of Sentential Truth Conditions .....	202
5.1 Deflationist Responses .....	207
5.1.1 The Challenge .....	208
5.1.2 The Objection .....	209
5.1.3 Attempted Accommodations .....	211
5.1.3.1 A Horwichian Accommodation .....	211
5.1.3.2 Field's Deflationary Strategy .....	214
5.2 Where Do We Go From Here? .....	227
5.3 Propositional Deflationism .....	231
Chapter 6: Concluding Remarks .....	237
References .....	242

## Chapter 0: Introduction

In this dissertation, I take a close look at the deflationary theory of truth, and deflationary semantics, generally. My thesis, baldly stated, is that, when we get clear on the central features of a deflationist account, we will find that, at present, the thorough-going deflationist picture championed by philosophers like Hartry Field and Paul Horwich is not yet at the stage where it can be adopted. If this is correct, then Field's *methodological deflationism* (1994a, p. 263), his view that we should assume deflationism as a working hypothesis, and that we should maintain it unless or until we find ourselves moved by some explanatory need to avail ourselves of more substantive machinery, should be resisted, or, at least, so I will argue. On the positive side, I will argue that there is an aspect of deflationism that seems unobjectionable, and I will show that the points in its favor do not support a thorough-going deflationist position.

In order to establish this thesis, we must first uncover the heart of deflationism, i.e., what it is for an account to be a *deflationist* account of truth. As we will see, there really has not been too much careful consideration of this question in the literature, the result being that there is more than one project that deserves to be called *the* deflationist theory of truth. So, in the course of pursuing our exploratory/evaluative project, I hope to clarify what makes a deflationary account of *X* a *deflationary* account, rather than an eliminative or reductive account.

Having made clear what makes an account deflationary, and, in particular, what features should be possessed by a deflationary theory of truth, I will make some problems

for the thorough-going deflationist, the deflationist about truth for all types of potential truth bearers. In so doing, we will uncover an interesting fact: there are deflationist programs in place that are pursuing radically different projects.

The contrast that I will make is between the Ayer- and Quine-inspired deflationism of Hartry Field, and the Wittgenstein-inspired deflationism of Paul Horwich, on the one side, and the more obviously Tarski-inspired deflationism of Scott Soames, on the other. According to Field, a theory of truth might just as well be called a theory of truth conditions—for him, to adopt a deflationist position on truth is to adopt a deflationist position on the explanation of truth conditions. Since, in the Frege-Russell tradition, it has been thought that there is an intimate connection between meaning and truth condition, such a view is deflationist about meaning, at least to the extent that it excludes truth conditions from playing a role in the theory thereof.

Soames' view of truth is considered deflationist as well, yet his deflationist theory of truth is hardly a deflationary theory of truth conditions: he takes very seriously the project of explaining how it comes to pass that a sentence, *s*, has the truth condition that it has. Furthermore, he takes meanings to be Russellian propositions, and, therefore, encoders of truth conditions. As we might put it, his view is deflationist about truth without being deflationist about semantics, so much so, in fact, that he isolates as centrally important the role of truth in the explanation of meaning and other semantic concepts. Hence, this version of the view is not the radical account floated by Field and Horwich; in fact, it is not even incompatible with a traditional (i.e., substantive) theory of meaning, though it is incompatible with the traditional correspondence theory of truth,

replete, as it is, with its ontology of facts, states of affairs, etc., and its ideology of proposition-to-fact correspondence relations.

To my knowledge, the question of how these two very different sorts of views could both be considered to issue in deflationary theories of truth has not yet been considered. In this dissertation I will address the question. As I will show, there is a difference between a deflationary theory of truth and a deflationary semantics, and it is this difference that distinguishes a view like Soames' from a view like Field's or Horwich's. If, as I will argue, the heart of deflationism resides in each, so to speak, then we will see that deflationism about truth might be a less contentious, more intuitive position than some have thought. However, if this is correct, then it will also follow that deflationism does not have the eliminative consequences that it is thought to have.

### **0.1 Deflationary Truth and Semantic Deflationism**

Deflationary semantic theories attempt to explain the semantic data by recourse to simple schemata, rather than by recourse to facts about the nature of meaning, truth, reference, or any of the other semantic properties that have made an appearance on the deflationist's chopping block. This has led to confusion, for it is not at all clear that the *traditional* questions—questions like What is truth? In virtue of what does a mark on a page or a mental state mean one thing rather than another?—can be answered by recourse to simple schemata such as the disquotational schema, (DS)

(DS) “p” is true iff p.

Hence, many philosophers—both inside and outside the philosophy of language—have been skeptical about the plausibility of the austere semantic picture being presented by

the deflationist, confused by the thought that the minimal resources the deflationist allows herself could suffice to explain the rich set of facts that the philosophers of language and mind are expected to tackle.

I think that at least some of the confusion about deflationism is legitimate, though much of it is illegitimate. Legitimate confusion stems from the fact that there is no consensus over what exactly a deflationist picture is to look like— even among deflationists. I believe that before we can even consider seriously the viability of a deflationist account, we must get clear on what should, and on what should not, count as a deflationary view. To this end, we will discuss the history of deflationism, distinguishing the negotiable from the non-negotiable features of the program. As I hope to show, there is a way to understand deflationism that distinguishes it from the inflationist (or, substantive) views against which it has been developed, but also from the eliminativist views with which it is sometimes confused.

Legitimate confusion also stems from the fact that there is an ambiguity to the appellation “deflationary theory of truth”, or, at any rate, so I would argue. As mentioned previously, there are at least two projects that a deflationist might be engaged in, one that focuses on the traditional question of the nature of truth, the other that places this question within a broader context and tries to provide a deflationary account of all of the semantic notions.<sup>1</sup> For purposes of taxonomy, we can distinguish between the two projects along the following dimensions: the first is largely a *metaphysical* project, whereas the second is largely a *metasemantical* project, though it entails, or near enough, conclusions like those reached in the metaphysical project. I do not believe that this

distinction has been made before, yet it is a very important one, for without it, as mentioned previously, we find that it is utterly baffling how both Hartry Field and Scott Soames could both count as deflationists, though their semantic pictures are in complete opposition. According to Field (1986, 1994a,b), a deflationist about truth is, first and foremost, a deflationist about *truth conditions*, and, in particular, about the explanation for how utterances and mental states get them. More to the point, he holds that we cannot explain how utterances and mental states get their truth conditions; per Field, an utterance (or, mental state) has the truth condition that it has by stipulation, more or less as a consequence of the way in which we have decided to use the word ‘true’.

Soames, as is well-known, accepts a very strong direct-reference picture of semantics, arguing (to the point of absurdity, some would say), that, even in the context of an attitude ascription, the reference of a ‘that’ clause is a Russellian proposition, i.e., an ordered n-tuple of objects and properties. Hence, Soames holds that we can explain how utterances get their truth conditions, and that the fact that an utterance has the truth condition that it has depends, in important ways, on facts about our language use.

I believe that the metaphysical/metasemantical ambiguity in the appellation “deflationary theory of truth” explains this seemingly odd circumstance. A metaphysical deflationist is deflationist—we will say what that means in due time, of course—about the nature of the property of truth. According to Soames, truth applies primarily to propositions. Hence, Soames is deflationist about what it is for a proposition to be true (or false). For this reason, we will call him, and, indeed, all metaphysical deflationists, *propositional deflationists*, since, as we are understanding the distinction, the

---

<sup>1</sup> I distinguish, as well, between a deflationist theory and a quietist theory, although I will not discuss the

metaphysical deflationist denies that deflationary claims about truth entail deflationary semantic claims, which means that she denies that linguistic and psychological objects—sentences, utterances, beliefs, and the like—are the primary bearers of truth. Soames, though he is deflationist about propositional truth, is not deflationist about what it takes for an utterance (or a speaker, depending on one's view) to express a proposition. Since he and all other propositionalists understand *what it is for an utterance to express the proposition that it does* to mean *what it is for an utterance to mean what it does*, Soames is not deflationist about meaning, and, therefore, though he endorses a deflationary theory of truth, he does not place this theory within the context of a deflationary semantics.

A semantic deflationist—Field is a paradigm case—is deflationist about more than just truth. A semantic deflationist adopts a deflationist attitude towards all of the central semantic concepts, *meaning* included. The semantic deflationist sees her account of truth as part of a larger picture within which she offers an alternative to the traditional (i.e., substantive) accounts of the notions of *meaning*, *content*, *reference*, and *truth*. For the majority of the dissertation, I will concern myself with this more comprehensive, thorough-going form of deflationism, returning to propositional deflationism only at the end, after we have made some trouble for the more thorough-going view, and when I draw some conclusions about what I think is central to the deflationary theory of truth.

## 0.2 Confusion Over the Deflationist's Claim

Aside from the confusion stemming from the fact that there are at least two different sorts of deflationist theories of truth, there is another confusion, largely borne of

---

alter in this dissertation. For a discussion of Quietism, see Wright, 1992.

ignorance of the nature of the deflationist project, which project is much different from the traditional project, the project of attempting to say what truth, reference, satisfaction, and allied notions like meaning and content *are*. In order to see this, let's consider, in broad terms, some of the "hard" questions asked in the philosophy of language, and in the relevant areas of the philosophy of mind, that have been thought to be central to our thinking about truth, and about semantics, generally, i.e., questions like (1) - (3).

- (1) In virtue of what does this brain state represent *snow*, rather than grass? (and what is it for one thing to represent another?)
- (2) In virtue of what is the following patch of ink—**snow is white**—true if and only if snow is actually white? (and what is it to say that it is true?) and
- (3) In virtue of what does my utterance of the sounds **Bill Clinton** refer to Bill Clinton (and what is for one thing to refer to another?)

For the naturalistically-minded philosopher, the project is in naturalistically-acceptable terms to explain and account for the occurrence of intentionality in the natural world. However, the urgency of the question should be obvious to anyone concerned with understanding language, thought and how they might be possible, no matter in which metaphysical camp she chooses to reside.

If the aforementioned questions are the ones that need answering, then, to all appearances, the thorough-going deflationist is lost,<sup>2</sup> for insofar as the deflationist holds that there is much less to meaning, truth and reference than people have thought, as we will see, she throws out the very tools that might be used to answer these questions.

---

<sup>2</sup> In what follows, I will drop the qualification 'thorough-going', except when it seems important to do otherwise. I will distinguish, then, between the philosopher whom I will call the deflationist—i.e., the thorough-going deflationist—and the philosopher whom I will call the propositional deflationist.

Although the deflationist will have trouble with (1) - (3), she is at home explaining facts like (4) - (6), which she takes to be analytic consequences of our use of the words 'mean'

(4) The (English) word 'snow' means *snow*

(5) The (English) sentence "snow is white" is true iff snow is white"

(6) The name 'Bill Clinton' refers (in English) to Bill Clinton

'true' and 'refers',<sup>3</sup> which use is governed by schemata such as (DS). However, (4) - (6) are not questions that have vexed philosophers for more than two millennia, and this is partly responsible for the confusion. The really challenging questions are, or, at any rate, appear to be, (1) - (3), for, though (4) - (6) may be analytic consequences of our use of the words 'mean', 'true' and 'refers', it is not an analytic consequence of anything that these putatively inert items—a patch of ink or a burst of air—have the semantic properties that we ascribe to them. Hence, that these physical objects have these semantic properties—if, indeed, they do—is not the sort of thing that can be explained by reference to the deflationist's schemata. If these putative facts need explaining, the explanations must cite very special features of us and them in terms of which it can be shown how intentionality emerges, and philosophers from Plato to Fodor have tried to say what those features are.

(1) - (6) capture meaning and reference facts,<sup>4</sup> but the issue regards truth as much as meaning. Consider for example the different sorts of explanation we would give for (7) and (8).

---

<sup>3</sup> Note that some deflationists—for example, Hartry Field—might prefer to speak only of idiolects, rather than languages like English. He would recast (4) - (6) in terms of a speaker's idiolect.

<sup>4</sup> I take facts about truth conditions to be meaning facts, since, on the traditional substantivist view, it is the responsibility of a theory of meaning to explain how sentences get their truth conditions.

(7) The (English) sentence—‘Snow is white’— is true

(8) The patch of ink—**Snow is white**—is true.

We might explain (7) by reference to an appropriate instance of (DS), together with the fact that snow *is* white,<sup>5</sup> but this would hardly be acceptable if what needed explaining was (8). What this means is that if (1) - (3) and (8) are the appropriate explananda for a philosophical account of meaning, reference, and truth, then a thorough-going deflationist account appears to be wildly untenable, and will not even begin to look adequate. It is because many philosophers take these to capture the central questions and facts that must be answered and explained that many philosophers have thought deflationism to be hopelessly and radically untenable. If the deflationists took (1) -(3) and (8) to capture the primary questions that needed answering, the central data that needed explaining, deflationism *would* be wildly untenable.

This explains the confusion over the likelihood of deflationist success: if the thought is that the deflationist will be able to explain everything the traditional philosopher—we can call the traditional philosopher the *substantivist*, which is less derogatory sounding, though no less accurate, than the usual appellation, *inflationist*—hopes to explain *in the way that the traditional philosopher*—the substantivist—*does*, then deflationism is destined to fail. However, the deflationist does not aim to answer the traditional questions, at least not in the way that the substantivist tries to answer them, and this is the first point at which the deflationist’s project departs from the norm.

If deflationism is to be at all tenable, the deflationist will have to reverse the usual order of explanation, claiming that, contrary to what we might have thought, (4) - (6) and

(7) furnish the central data (or readings), and the other readings are, in some sense, derivative on those. The deflationist, whether she takes propositions, utterances, statements, or whatever to be the primary bearers of semantic properties, divides explananda. On the one hand, there are questions about the truth conditions, meanings, truth values and references of specific sentences, words, and phrases, questions that can be asked and answered only from *within* a language, i.e., when considering *interpreted entities*. These are questions that can be answered, to a greater or lesser extent, by all competent speakers of the language, in virtue of their linguistic competence. They are ordinary questions, i.e., about whether or not a sentence is true; about what an expression refers to; about whether one word means the same as another; etc. These are the questions that the deflationist attempts to explain by recourse to the minimal machinery she avails herself of, such as the schema, (DS), the conceptual roles of expressions and sentences of the language, and non-linguistic facts, such as the color of snow. Her claim is that *these* sorts of facts—internal facts, as we might call them—can be explained without wheeling out any heavy-duty metaphysical machinery.<sup>6</sup>

These internal questions, for the most part, are not the ones that have exercised philosophers, and, as the deflationist is quick to point out, they do not really support the search for an explanatory *theory* of anything. So what about the traditional questions, the hard ones that philosophers do worry about? Here, the deflationist cannot avail herself of

---

<sup>5</sup> Alternatively, if we take the sentence to be fully interpreted, as Field (1994a,b) does, we might explain it by just talking about the color of snow.

<sup>6</sup> The parallel with Carnap's internal/external distinction should not be missed. I have come to believe that this is a distinction that should be revived, in limited form, by the deflationist. My reasons for thinking this come from my own thinking about how best to understand deflationism, and from reading some of the work done by Huw Price. The picture that Price paints is different in important ways from the one that I am sketching, though mine is very clearly inspired by his.

her schemata or conceptual roles, for they will do no good. However, rather than advert to underlying facts, causal connections, congruence relations, or any of the ontological or ideological constructs that philosophers have developed to try to explain meaning, truth and reference, the deflationist adverts to facts about the *function* of our semantic vocabulary. Rather than explaining what truth or reference *are*, i.e., in reductive terms, they explain what ‘true’ and ‘refers’ *do*, i.e., why it is so important for our language to have predicates that play the role that is played by these expressions, and what role truth or content ascriptions play in our practices.<sup>7</sup> To take an example, something like (9) will be read by the deflationist as (10),

(9) This patch of ink—**snow**—means *snow*

(10) This patch of ink—**snow**—will be read (interpreted) by competent speakers of English as the word, “snow”, which means *snow*,

where the semantic properties are not ascribed directly to the *physical* objects,<sup>8</sup> which objects are ascribed the property of being read or interpreted by speakers of English in a particular way, but to the *semantic* objects, the expressions of English. To be sure, the deflationist has to explain the phenomenon cited in (10), but doing so will require providing a much different sort of account from the reductive ones that we have come to

---

<sup>7</sup> This is something of an idealization. As I will describe it, this is what the deflationist must do, because this is the key to her non-reductive explanatory strategy. The problem is that most deflationists have not exactly sketched out the linguistic functions of *refers* or *means that*. A noted exception is Bob Brandom (1976, 1984, 1997, and 1998). Although I find Brandom’s work interesting, I will have to leave a discussion of it to another time.

<sup>8</sup> It should be clear, then, that I think that it is a mistake on Horwich’s part (1997, 1998) to claim, as he does (e.g., 1997, p. 99), that he is trying to answer the question “How is it possible for a word—a mere mark or sound—to represent (refer to, be about) some specific aspect of the world?”. This is not something that the deflationist can explain, and it better not be something that she thinks needs to be explained, so I will take Horwich to have confused the deflationist’s project with the project that it aims to replace.

expect; instead, it will involve explaining the practice of ascribing semantic properties, and the function served by such ascriptions.

This project, too, has been confused, and even by some deflationists—see my discussion of Horwich, in chapter 4—because the old explanatory project is deeply entrenched in our philosophical practices. To avoid confusion, let's consider for a moment the basic format of the substantivist's account. Confronted with putative facts like (8) and (9), the substantivist aims to give *elucidations*, where the elucidation of a property,  $\Phi$ , is an attempt to exhibit non-analytic connections between  $\Phi$  and other properties,  $\Psi$ ; i.e., connections expressible as (E).

(E) For all  $x$ ,  $x$  is  $\Phi$  iff  $x$  is  $\Psi$  (or, if; or, only if).

That is, the substantivist attempts to analyze or reduce the notions of meaning, truth and reference, to expose their natures and show us, in non-semantic terms, what they really are.

By contrast, the deflationist is not interested in this sort of elucidation. The deflationist program (*pace* Horwich) does not involve an attempt to give a reductive analysis of anything, or, at least, of the central semantic concepts that have been the targets of traditional analyses.<sup>9</sup> Instead, the deflationist aims to explain our *talk* of

---

<sup>9</sup> The qualification is important because it is not clear that the deflationist does not want to jettison elucidations altogether. One way of viewing Horwich's program, as that has developed, is as follows: Although he merely provides an implicit definition of truth (i.e., an *axiomatization* of that notion), he provides an explicit definition of what it is to grasp the concept of truth, which is to say that he gives an elucidation of that notion. As he sees it, to grasp the concept of truth is to take instances of (DS) to be conceptually fundamental.

meaning, truth and reference, and to say what a person must “know”<sup>10</sup> in order to be competent with the relevant concepts.

Why do the deflationists offer alternative sorts of analyses? Although not all deflationists make explicit a general motivation for adopting the position and offering an alternative to the reductive analyses favored by the substantivist, and although what follows may not even hold for all deflationists, it seems that we can explain a central motivation driving at least some, if not all, versions of deflationism. The deflationist believes that the substantivist has made a faulty assumption, the assumption that reductive analyses are appropriate for the denizens of all domains. As the deflationist sees it, reductive analyses are appropriate only when the reducing theory and the to-be-reduced theory play the same linguistic function (e.g., fact stating). If, as the deflationist claims is the case with truth, the explanatory role of the property is exhausted by the logical or expressive function of the concept—in this case, for English, ‘is true’, the truth predicate—then it is inappropriate to expect a reduction to, say, the language of physics, for, *ex hypothesi*, the concepts of physics and the concept of truth play very different linguistic roles. So, the deflationist holds that whether or not we should expect a reductive analysis depends on what the candidate for reduction does; that is, it depends on its explanatory role. Hence, in answer to our question at the top of the paragraph, we can say that, in the case of truth, the deflationist offers an alternative to the reductive analysis because she believes that the function of that property is exhausted by the function of the concept of truth, and that we can provide a complete account of the function of the concept of truth by laying out its logical (or, linguistic; or, expressive)

---

<sup>10</sup> I put scare quotes around ‘know’ to signal that I do not intend it in the sense of propositional

role.

Of course both the substantivist and the deflationist have as their goals the explanation of certain phenomena, but the phenomena thought to need explaining are not the same. The substantivist assumes, or believes, that the phenomena include such things as that certain sentences—or, for example, marks on a page—have the truth conditions they have; that certain words—or, for example, vibrations of air—represent what they do; and that certain thoughts—or, for example, brain states—are either true or false. By contrast, the deflationist does not assume that our talk of truth, meaning, reference, and the like involves the ascription of *genuine* or substantive properties, to marks on a page, vibrations of air, or states of a brain, for the deflationist believes that we can explain everything that needs explaining by illuminating the *function* of our semantic talk. In particular, then, the deflationist<sup>11</sup> does not assume that in ascribing meaning, truth, reference, and the like, we ascribe substantive ( e.g., causal-explanatory) properties.

### 0.3 My Thesis

The purpose of these last sections has been to take note of some confusions that have arisen with respect to deflationism, and to set out, in sketch form, a characterization of what I take to constitute the central features of a deflationary view. I believe that these features show deflationism about truth to bear important relations to views like non-cognitivism in ethics, and to bear important dissimilarities to certain eliminativist views with which it seems to have been confused. I will develop further the picture of the deflationist strategy that I began above, as the dissertation unfolds.

---

knowledge.

We have arrived at the topic of this dissertation. Although I find deflationism attractive, and although I believe that one form of deflationism—propositional deflationism; metaphysical deflationism—is both philosophically<sup>12</sup> unobjectionable and more or less correct, I am not yet convinced that a thorough-going deflationism—deflationism for all putative truth bearers—is acceptable. I have been led to believe this because I have come to believe that it is much harder than many have thought for the thorough-going deflationist to accommodate our intuitions about the relation between meaning and truth, and it is not yet clear to me how she will go about doing this. I will bring this fact out in chapter 5. I am also skeptical about the ability of idiolect disquotationalism to issue in a notion of truth adequate to perform the function that we expect a truth predicate to perform, as I will mention in chapter 2.

The remainder of my dissertation is divided into 4 chapters. In chapter 1, I survey the history of deflationism, determining which of the theses that have been put forward are relevant, which are irrelevant, and how we ought to understand the variety of deflationism.

In chapter 2, I turn to contemporary accounts—primarily the accounts developed by Hartry Field, Paul Horwich, Stephen Leeds, and Vann McGee—and I survey and critique the two most promising-looking versions of deflationism, Horwich’s property deflationism, and the idiolect deflationism developed by Hartry Field. In the end, I conclude that *property deflationism* is less contentious than idiolect disquotationalism, and, for this reason, it is the version that I assume in the remainder of the dissertation

---

<sup>11</sup> Recall that I use ‘deflationist’ here when I intend to speak about the thorough-going deflationist, i.e., the deflationist about semantics, generally.

<sup>12</sup> I add the qualifier because of doubts that have been registered regarding its technical correctness.

(though I spend a fair amount of time talking about idiolect disquotationalism). I try to motivate property deflationism because it is a liberal form of deflationism. Since I want to give the deflationist the best run possible, it is best to stick with the version that has the fewest extraneous theoretical commitments, and I believe that property deflationism is such a view.

In chapter 3, I consider claims about the explanatory role of truth and I attempt to show that one form of argument—an argument from the explanation of behavior, in which it is claimed truth plays a causal-explanatory role—can be defused and that, consequently, one class of argument that has been thought to threaten the deflationist is a non-starter. However, I will note that there are certain putative facts about truth that do not seem explicable on the deflationist's view.

In chapter 4, I consider what the deflationist about truth has to say about meaning, concentrating, in particular, on the accounts proffered by Paul Horwich and Field. As I show, Horwich's argument—to the effect that the deflationist can explain the so-called meaning-to-truth conditionals, which conditionals she takes to underwrite the possibility of representation—fails, and I show that, given the meaning-to-truth conditionals, it can be established that a thorough-going deflationist must be a deflationist about meaning as well as about truth. This result, though quite general, shows Horwich's attempt to combine a deflationary account of truth with a use-theoretic account of meaning to be misguided.

The use-independence of truth acknowledges that the truth (or truth condition) of an utterance does not depend on its having the meaning that it has. Given the argument of chapter 4, to the effect that no thorough-going deflationist can countenance a non-

deflationary theory of meaning, this is a good thing, for the thorough-going deflationist does not allow herself the semantic resources needed to explain the contingency of truth and truth conditions on meaning. As Field (1994a) notes, this appears to lead to problems, for it suggests that, say, even if we had used ‘snow’ to mean what we now use ‘grass’ to mean, ‘snow is white’ would have been true. This has been acknowledged by all—deflationists and substantivists alike—to be a counter-intuitive result of deflationism. I do not find the claim that it is counter-intuitive to be terribly interesting, or, indeed, particularly threatening to the deflationist, *provided she can explain away our intuitions*. As I show in chapter 5, the deflationist strategies that have been used (and those that might be used) do not succeed in explaining away our intuition of the contingency of sentential truth conditions. If we can establish that the intuitions need to be explained, and we can establish that the most plausible and straightforward explanation of the intuitions is that the truth conditions of an utterance depend on that utterance’s meaning what it does, then we will have created a relatively serious problem for the deflationist. I think that we can show that the intuition needs explaining, and, therefore, I think that we can show that the deflationist is in trouble.

In the concluding chapter, I discuss and attempt to motivate what I call propositional deflationism, a version of deflationism restricted to propositions only. I believe that propositional deflationism is an intuitive and well-motivated position, and that it is what is suggested by the desire to avoid the ontology and ideology of at least the traditional correspondence theory of truth. I support propositional deflationism by showing that the deflationist about truth need not be a semantic deflationist; that the explanatory role of truth does not support semantic deflationism, which position could

only be supported by the provision of something like a deflationary account of meaning; and that some of the conceptual data that has been thought to be central to deflationism do not allow us to select from between the two views. Hence, my arguments present a challenge to the deflationist: Show that the explanatory role of meaning can be captured in an account of the concept thereof, N.B., where such an account would not include a reductive analysis of the property of meaning, or give up semantic deflationism, and be happy with propositional deflationism.

This is my plan. Here is the dissertation.

### **Chapter 1: The History of Deflationism**

Although I agree that there is no deflationist theory of truth, *per se*, I do not believe that the various forms of deflationary truth are unified only in their rejection of all attempts to say anything of interest about truth and truth conditions, as some have thought. In fact, as I will try to show, there is a central thesis that every position that has been thought to qualify as deflationist must adopt. The feature that has been thought to unify deflationist theories is acceptance of the claim that *truth*<sup>13</sup> is not the sort of property for which we should expect a reductive analysis.<sup>14</sup> Although this does unify all of the putatively deflationist theories, it is too rough-hewn, for it does not allow us to distinguish deflationist from eliminativist or from inflationary primitivist theories.

As I would like to characterize the central thesis of deflationism, it involves a certain *explanatory* claim, namely, the claim that *truth* enjoys no reductive analysis *because* its explanatory role does not support a reductive analysis. On this view, we

---

<sup>13</sup> In what follows, I use italics to indicate properties and boldface to indicate concepts.

<sup>14</sup> This claim can be recast so as to be acceptable to the nominalist.

should expect a reductive analysis of a concept, only when the explanatory role of the property picked out by the predicate calls for one. So, to take an obvious (maybe the only) example, we expect reductive analyses of concepts that pick out properties that are used to categorize and explain certain events on the basis of characteristics shared by the entities that fall under them. Since *truth* is no such property, the deflationist claims, it is not a proper candidate for reduction. I take this to be a central thesis of deflationism, one that we will arrive at once we have traced the recent history of deflationism and have separated the wheat from the chaff.<sup>15</sup>

### 1.1 A Brief History of Deflationism

According to Dummett (1959), the history of deflationism begins with Frege, who (1918, pp. 354-5; reprinted in Salmon and Soames, 1988, pp. 36-37) points out that we cannot recognize a property of a thing without at the same time finding the thought *this thing has this property* to be true. So with every property of a thing there is tied up a property of a thought, namely truth. It is also worth noticing that the sentence ‘I smell the scent of violets’ has just the same content as the sentence ‘It is true that I smell the scent of violets.’ So it seems, then, that nothing is added to the thought by my ascribing to it the property of truth. And yet is it not a great result when the scientist after much hesitation and laborious research can finally say ‘My conjecture is true’? The meaning of the word true seems to be altogether

---

<sup>15</sup> I should note that, in claiming this to be a central feature of deflationism, I do not mean to say that there is some one view that is intended by *deflationism*, nor do I mean that we cannot talk about views having different deflationary features. However, as I hinted at above, I believe that one of the problems with the debate is that it is simply not clear what makes a view deflationary. My introduction of the explanatory requirement is an attempt to say what I think (is a part of what) really distinguishes deflationary theories from other sorts of theories.

*sui generis*. May we not be dealing here with something which cannot be called a property in the ordinary sense at all?

In Frege's quote, we can discern two important claims, one that divides the contemporary deflationist views we will be discussing, and one that unifies them. The divisive point regards truth bearers (or *primary* truth bearers, for Frege does not deny that we can attribute the property of truth to sentences)—per Frege, truth is ascribed primarily to thoughts (or *propositions*), which are the contents of sentences and attitudes (e.g., belief, conjecture, etc.).

The unifying point regards Frege's important distinction between two *uses* of 'true'—*free standing* uses and *blind ascriptions*, as they have come to be called. Free standing uses, i.e., a use of the form 'it is true that p' or 'that p is true'—a use in which truth is ascribed directly to a 'that'-clause—appear to be *transparent*: as Frege would put it, in asserting, believing or conjecturing **that p is true**<sup>16</sup> one asserts, believes or conjectures nothing more than one would if one asserted, believed or conjectured **that p**, and (presumably, for one who grasped the concept of truth), vice versa. This is what has been called (e.g., by Blackburn, 1984, p. 227) the *transparency thesis*, and, it is claimed, it follows from it that **It is true that p** and **p** are trivially equivalent,<sup>17</sup> from which follows the equivalence schema, (EO).

(EO) It is true that p iff p<sup>18,19</sup>

<sup>16</sup> In the following, I use boldface for concepts and propositions and italics for properties.

<sup>17</sup> I use this neologism because there are various (and incompatible) claims about what follows from the transparency thesis.

<sup>18</sup> Actually, this claim must have to be restricted, in light of the possibility of truth-value gaps, such as might arise in regards to paradoxical sentences (or the propositions they express)

As Blackburn sets it up, according to the transparency thesis, anyone who believes *p* also believes **it is true that *p***.

In the case of a free-standing use of ‘true’, as in (11)

(11) It is true that I loath the scent of violets,

nothing is added to the thought by the addition of the truth ascription. In “blind ascription” of ‘true’, such as is found in (12) or (13),

(12) My conjecture is true

(13) Everything Clinton says is true

something is added, according to Frege—indeed, the truth predicate is ineliminable (or near enough), given the way our language is set up.<sup>20</sup> In this way, the second sort of use provides us with a reason for having the truth predicate: the truth predicate allows us to do things that we could not otherwise do, or, at least, that we could not as cheaply do otherwise, i.e., it adds to the *expressibility* of the language, allowing us to make generalizations that we could not otherwise make.

Hence, in Frege’s quote, we discern two central features of contemporary deflationist views. First, from Frege’s observations about the first type of use, we can get the beginnings of an account of what, on the deflationist view, truth *is*; second, from Frege’s observations about the second type of use, we get the beginnings of an account of what, on the deflationist view, truth *does*, i.e., what it is good for. It is a short step from

---

<sup>19</sup> We should also note that (EO) differs from (E) in that it employs a truth operator, rather than a truth predicate. We should note that we could recast Blackburn’s claim using the truth predicate.

<sup>20</sup> I say this because, as Horwich (1998, pp. 85-86) points out, the concept of truth is not needed for the purposes of generalizing; our language *could* have a substitutional quantifier. However, given the extra semantic and syntactic machinery that would be needed, there is good reason for having the truth predicate: it is, as Horwich puts it, a ‘cheap’ way of obtaining the generalizations.

Frege's thoughts on truth to the contentions made by Field (1994), Horwich (1990, 1997, 1998), and Leeds (1978, 1995) that (14) and (15) are central to the deflationist program.

(14) 'It is true that p' and 'p' are trivially equivalent.<sup>21</sup>

(15) The primary function of the truth predicate is as a device of generalization.

A consideration of (14) and (15) is of central importance to understanding the history of deflationism, for the varieties of deflationism have developed out of different ways of accommodating them. Hence, we have the following historical "progression":

(C1) The Classical Redundancy theory: 'true' expresses no property; **it is true that p** and **p** "say" the same thing

(C2) The Tarskian theory: an adequate definition of truth must entail all instances of (DS).<sup>22</sup>

(C3) The Quinian theory: to say that 'snow is white' is to call snow white; truth is just disquotation.

(C4) Fieldian Deflationism: *u* is true is cognitively equivalent to *u* relative to the utterance of *u*;

(C5) Horwichian Deflationism: the disquotational schema, "p" is true iff p, is conceptually fundamental, meaning that we accept its instances in the absence of supporting argument.

It is not my purpose to give an exhaustive account of the history of deflationism; rather, I want to say what is necessary about the positions and their motivations in order

---

<sup>21</sup> They differ on matters of truth bearer, as we will discuss. Horwich takes the truth bearers to be propositions, while Leeds and Field take them to be utterances. Furthermore, we will discuss later how 'trivially equivalent' is to be understood.

<sup>22</sup> The reader might be surprised by the order, here: shouldn't Tarski's theory go before Quine's? The reason for setting it out in this order will be obvious, in what follows.

for us to understand the twists and turns of the contemporary debate, i.e., the positions championed by, among others, Field and Horwich.

## 1.2 The Redundancy Theory

The redundancy theory is motivated by a consideration of the first type of use of ‘true’: If there is a difference in content between the proposition expressed by an utterance of (16) and the proposition expressed by an utterance of (11) then why is it that, (16) I loathe the scent of violets,

(11) It is true that I loathe the scent of violets,

for one who grasps the notion of truth, asserting or believing the one involves nothing more than is involved in asserting or believing the other? If ‘true’ expresses a genuine property, what is that property like such that the aforementioned putative datum can be accounted for?

It is skepticism about there being any such property that motivates the redundancy theory: the best explanation for the aforementioned (putative) fact is that ‘true’ expresses no property at all. On this view, at least as developed by Ramsey (1927) and by Strawson (1950),<sup>23</sup> apparent predications of truth have no descriptive function; an utterance of (11) is just a long-winded way of saying what is said in an utterance of (16).<sup>24</sup> Hence, the Redundancy theorist concludes, there is no property of truth, and, appearances to the contrary notwithstanding, in the making of a truth ascription, we do not describe the

---

<sup>23</sup> I do not mean to lump together Ramsey and Strawson as supporters of the same view. However, to the extent that their views intersect, they intersect here.

<sup>24</sup> To say that truth ascriptions have no descriptive function is not to deny them any function at all. For example, Strawson assimilates truth ascriptions to performative utterances, claiming that when we appear to ascribe truth to a proposition, what we are doing is endorsing the proposition.

propositions to which we appear to ascribe truth; *a fortiori*, we do not describe them as being a particular way (namely, as being true).

There are two things to note. First, we should note what is wrong with the view. Second, we should note that, though, evidently, Frege endorsed a sort of redundancy account, it is not clear that the conclusion of his most ingenious deflationist argument is the redundancy of truth ascriptions.<sup>25</sup> In fact, I think, we can learn a tremendous amount about contemporary deflationism, by paying careful attention to this argument, which we will turn to in the next section.

Ramsey's claim, then, in saying that truth ascriptions are redundant, is that 'is true' is *vacuous*, that it says nothing at all: "It is true that Caesar was murdered' says nothing more than 'Caesar was murdered'." (1927, pp. 16-17) Hence, Ramsey endorses schema (DR), where ' $\leftrightarrow$ ' is interpreted as *is analytically equivalent to, is synonymous with, or means the same thing as*.

(DR) The proposition that p is true  $\leftrightarrow$  p.

As noted previously, there are (at least) two uses of the truth predicate. While this account might be acceptable for free-standing ascriptions, it does not seem so for blind ascriptions. If, for example, you believe Clinton to be a generally honest guy then, even if you do not know what he has said recently, you might utter something like (17).

(17) Everything Clinton says is true.

---

<sup>25</sup> Frege's primary reasons for adopting a redundancy account, aside from those mentioned here, are bound up with his concept/object distinction. A thorough discussion of these epicycles would take us pretty far afield, and for little gain, as far as I can tell.

*Ex hypothesi*, you do not know what Clinton has said; hence, holding that you have said just what Clinton did is dubious. Although Ramsey's account was adequate for our first type of use, it appears to founder on our second type of use.

Ramsey was aware of this consequence of his view, and he did an interesting thing: rather than ditch the redundancy account, he offered a different account for blind ascriptions, one on which the truth predicate is not redundant in the way it is in the first account. However, as Davidson notes (1990, 282-283), the truth predicate is still redundant in a sense, for everything that we can say with it (in principle) can be said without it. Ramsey's account of (17) would be something like (18), where the variable, P, occupies the position of a sentence,

(18)  $(\forall P)$  (Clinton asserted that  $P \rightarrow P$ )

which is to say, in particular, that an instance of (18) is formed by replacing P with a sentence, rather than with the name of a proposition.<sup>26</sup>

There are reasons for thinking that (18) is problematic,<sup>27</sup> for there are reasons for thinking that the Redundancy theorist cannot be comfortable with either a substitutional or an objectual reading of the quantifier. Problems with the substitutional interpretation of the quantifier come in when we consider propositions that are expressed by different sentences at different times. Richard (1990, Pp. 77-78) provides an argument for a related claim that goes something like this.<sup>28</sup> Consider a sentence like (19), where the quantifier

---

<sup>26</sup> Why not think that (2) should be recast as (2a), which is a more ordinary way of understanding (1), (2a)  $(\forall p)$  (Clinton asserted  $p \rightarrow p$  is true) that involves first-order quantification over propositions? The answer comes from Davidson's point: Ramsey is still offering a redundancy theory of sorts, and, as Soames (1997, p.6) notes, "[{(2a)}] contains a truth predicate [so] redundancy theorists cannot make use of it."

<sup>27</sup> Cf. Richard, 1990.

<sup>28</sup> I have simplified the argument so as to include only the essential parts.

(19) Ken Starr now believes everything he previously believed  
is understood substitutionally. If it is understood substitutionally, then if it is true, so is  
(20). However, it is obvious that (19) can be true though (20) is not, for what Ken Starr  
(20) At some earlier time Ken Starr believed  $P$  & Ken Starr now believes  $P$   
might have believed earlier is (21), for which (19), but not (20), is true.

(21) Today is January 15, 1998

Since, the truth of (19) requires the truth of (20) if the quantifier is understood substitutionally, and since (19) can be true though (20) is not, we should not take it that Ramsey's quantifier is understood substitutionally.<sup>29</sup> So we should take it that the quantifier in (18) is objectually quantifying over propositions.

The problem with this option is that it carries with it a commitment to propositions, and many who have adopted the redundancy theory—Ramsey and Strawson, in particular, but also Dorothy Grover—<sup>30</sup> have also been interested in eliminating propositions from their account. More important for our purposes than the historical point, if you countenance propositions, it is hard to see how you could deny that there is a property shared by all and only the true ones,<sup>31</sup> for if you admit that there are propositions, it is easy to define the property of truth as in (PT), where the quantifier (PT)  $(\forall x) [(x \text{ is true}) \leftrightarrow \exists P (x = \text{the proposition that } P \ \& \ P)]$  is understood substitutionally. This is a deflationist account of the property of truth, as that will be explained later, but if we can have as range of “ $x$ ” and “ $P$ ” the class of

<sup>29</sup> There are other problems with the claim that have been registered, but I will not address those here.

<sup>30</sup> Cf. Soames, 1997, p. 9, for a version of this argument.

<sup>31</sup> This point was made to me by Stephen Schiffer, and I think it is a good one. We should note, of course, that in saying this, we are not saying that if you admit propositions, you are committed to the claim that

propositions, we have a trivial definition of propositional truth.<sup>32</sup> Hence, even for the redundancy theorist who is not afraid of propositions, it looks as if it will be hard to give an objectual reading of the quantifier while maintaining that the truth predicate expresses no property. Although Ramsey's account is flawed in important ways, it makes an important contribution to deflationist thinking: Ramsey sharpened Frege's claim of *content redundancy*, according to which 'the proposition *that p* is true' is cognitively equivalent to '*p*'. This position is endorsed by contemporary prosententialists (e.g., Brandom, Grover), and, in a qualified form, by disquotationalists (e.g., Field) and Minimalists (e.g., Horwich).

### 1.2.1 Ayer's Redundancy Theory

There are three other versions of the redundancy theory to discuss: Ayer's, Strawson's and Grover, Camp and Belnap's. Ayer (1936) adopts what is, effectively, a Ramseyan account of truth: according to Ayer (1936, pp. 88-89),

[I]n all sentences of the form "*p* is true," the phrase "is true" is logically superfluous... Thus, to say that a proposition is true is just to assert it and to say that it is false is just to assert its contradictory. And this indicates that the terms "true" and "false" connote nothing... "true" does not stand for a genuine quality or relation.

What distinguishes Ayer's view from Ramsey's view is his theory of meaning, (i.e., of propositions, or of truth conditions). Ramsey assumes that truth applied to Fregean

---

there is a substantive property of truth possessed by all and only the true ones. An explanation of these vague remarks will have to await our discussion of the property of truth.

propositions, which encapsulate truth conditions. Hence, although Ramsey offers an eliminativist account of truth—he denies that ‘true’ expresses a property, after all—there is an inflationist aspect to his view, for he has an inflationist view of what it is for an utterance or belief to express a proposition. Although Field (1986) takes this as evidence that Ramsey does not offer a deflationary theory of truth at all, I think that we should say that Ramsey attempts to combine a deflationary (or eliminativist) view of truth with an inflationary view of meaning.<sup>33</sup>

Ayer, on the other hand, has a considerably more deflated view of propositions. According to Ayer (1936, p. 88), a proposition is “a class of sentences which have the same intentional significance for anyone that understands them.” This is where the *language independence* of truth conditions, so important to contemporary deflationists, comes from. For Ayer, the truth condition for a sentence, *s*, is an equivalence class closed under, as we might call it, *person-relative intentional significance*. What this means is that to claim that ‘*s*’ expresses the proposition that *s* is to say something necessary—*s* could not fail to be in an equivalence class that contained it—and it is to say something that is true independently of how we use the language. Hence, aside from

---

<sup>32</sup> Although this formulation is Soames’, we should add a few things. First, it is inconsistent, owing to the paradoxes. If our interest was in developing this as an adequate analysis of truth, we would have to say more about how to handle those. However, for present purposes, (5) will do.

<sup>33</sup> In saying this, I do not mean to deny Field’s claim that Ramsey has all of the resources to provide an inflationary account of the property of truth; I think that Field is right about this. What I object to in Field’s presentation is his (intentional) conflation of the question of what it is for an utterance to have the truth conditions that it has and the question of what it is for it to be true or false. Although I think these are closely related, I think that they are different questions, and that their close relation is a consequence of the close relation between (one aspect of) meaning and truth. I will discuss this more, when we turn to Field’s view.

deflating (or, as in Ramsey's case, eliminating) the property of truth, Ayer deflates the property of having truth conditions.<sup>34</sup>

I want to point out that, although Ayer discusses deflationary truth just after his discussion of deflationary propositions, and though it would be ridiculous to view the facets of his semantic theory as unconnected, he does not take the deflationary account of truth and the deflationary account of propositions to be identical, as does Field (1986, p. 55), when he (i.e., Field) says “as I understand a theory of truth, it could equally be called a theory of truth-*conditions*: the question that forms the subject of a theory of truth is substantially equivalent to the question of what it is for an utterance to have a given state of truth conditions.” We will discuss Field's position later, but it is important to note that Ayer's primary concern regarding truth (as opposed to propositions, which, for Field, means truth conditions) appears to be with the question of the propositional content added by 'true' to what is expressed by sentences containing it. What he says on this matter is that 'true' and 'false' mark sentences as assertions and denials, respectively. Hence, he, too, denies that attributions of truth are *descriptive*; rather he holds, they indicate the performance of a particular propositional speech act.

### 1.2.2 Strawson's Redundancy Account

Strawson agrees that 'is true' is not used to describe a proposition. He held<sup>35</sup> that it is used to perform some other sort of speech act, such as agreeing with a speaker, conceding a point, or what have you. Here's Strawson (1954, p. 271, 263): “The phrase

---

<sup>34</sup> In this discussion, I have followed Field (1986, pp. 56-61) and Ayer (1936)..

<sup>35</sup> He has since dropped the view, for reasons discussed below.

‘is true’ never has a statement making role... Truth is not a property of symbols; for it is not a property.”

Strawson has abandoned this view, and I imagine (or, hope) that this is because of his recognition of its two most devastating criticisms. First, Strawson’s view cannot deal with many contexts in which ‘true’ occurs, such as conditionals and generalizations, for example, (22). Second, and, I think, far more important, Strawson needs a premise

(22) You can only know what is true

to get from the function he isolates to the no-property claim, and the premise he needs he cannot have. Suppose that, after a long argument (which I have decided to let you win, though I still believe that I may be right), I say “OK; what you say is true.” Strawson is no doubt right that I am using ‘true’ in order to concede the point to you; however, what looks plausible is that I do so *by* ascribing truth to what you said—in describing what you said as true, I concede the point. Hence, we see that simply by identifying a (or even the) use of the truth predicate that does not require us to posit a substantive property of truth, one cannot automatically conclude that there is no property of truth. What is needed is an incompatibility argument: it must be shown that the claim that ‘truth’ plays this (even: and only this) role is incompatible with the claim that it expresses a property, *truth*. However, this incompatibility cannot be made out, for it is compatible with Strawson’s claim about the function of the truth predicate that it (or, the concept of truth) expresses a *thin* property (i.e., in a way to be discussed later).

### 1.2.3 Grover, Camp and Belnap

Grover, Camp and Belnap (1975) also argue that ‘is true’ is not a predicate, and that *truth* is not a property. They isolate a class of *pro* expressions, expressions with anaphoric functions. The claim that ‘true’-containing sentences: ‘that is true’, ‘it is true that snow is white’, etc. are pro-sentences, equivalent to sentences like “Snow is white”, etc. Furthermore, for blind uses, they assimilate the use of ‘true’-containing pro-sentences to other so-called pro-forms. Hence, (22) is understood as (23),

(22) Everything Bill said is true

(23) For every proposition, if Bill said it is true, it is true

where the quantification is understood substitutionally. What is important is not that ‘true’ be eliminated, but that it not be treated as a *separable* predicate: ‘true’-containing pro-sentences are treated as semantically unanalyzed units, and ‘true’ is not a *genuine* property-expressing predicate, for it is not a genuine predicate at all. Hence, although ‘it is true’, or ‘that is true’ cannot be eliminated from the language, ‘true’, *qua* separable predicate, can be. Hence, theirs is not exactly a redundancy view, but it is close, for ‘true’-containing pro-sentences have only an anaphoric function.

Although Grover, Camp and Belnap offer an interesting, non-descriptive (i.e., non-classificatory) account of truth, ultimately, their account fails. Perhaps the most devastating criticism that can be made of this view comes out when we let Grover, Camp and Belnap theorize themselves into a corner. The first question one should ask is why it is that ‘true’ appears to be a separable predicate, though it is not. Per Grover, Camp and Belnap (1975, pp. 99-100), the answer comes in two steps, corresponding to the two uses mentioned above. First, they claim that we have reason for wanting non-atomic pro-forms so that we can modify tense and mood, and so that we can negate, i.e., so that we

can say ‘that was true’, ‘that might be true’ and ‘that is not true’. Second, they claim that the subject-predicate form is so widely used that it makes sense that we should have a way of collapsing the conditional, (24) into (25).

(24) For every proposition, if Bill said it is true then it is true.

(25) Everything Bill said is true.

They claim (1975, p. 86), then, that they can handle all of the modified uses of ‘it is true’, simply by postulating a limited number of deep structure interpretations, namely, ‘it-is-true’, ‘it-might-be-true’, ‘it-was-true’, ‘it-will-be-true’, and ‘it-is-not-true’. Hence, a past tense version of (24) would be (26). The problem is obvious: ‘true’ appears in the

(26) For every proposition, if Bill said that it was true, then it-was-true that it is true.

deep-structure sentential operators that are posited. If they can be analyzed away, why posit them in the first place? If they cannot, how are the prosententialists going to avoid a commitment to the property of truth?

It should be objected that I am overlooking the obvious role of the hyphens. Of course it *appears* as if the prosententialists are in trouble, but that is only because we are neglecting the fact that ‘true’ in ‘it-was-true’ is not its own predicate, but is part of the operator, fused in Quinian fashion. I am certain that this is how the prosententialists want us to understand the occurrence of ‘true’ in the aforementioned fused operators. However, if this is what they mean, then they are in trouble. They must explain, without employing the concept of truth, what ‘it-was-true’ means, and this they do not seem to be able to do.

### 1.3 Frege’s Anti-Correspondence Argument

We return to Frege's anti-correspondence argument. Frege explicitly accepts the redundancy claim, and the following selection, also found in "The Thought", provides an argument for it. We return to the challenge previously floated: if there is an actual difference in content between what is said or thought when one says or thinks *that p*, as opposed to *it is true that p*, why is it not the case that everyone who understands the notion of truth and who  $\Phi$ s the first, also  $\Phi$ s the second, for any propositional attitude,  $\Phi$ ? Here's the argument (Salmon and Soames, 1988, p. 35)

But could we not maintain that there is truth when there is correspondence in a certain respect? But which respect? For in that case what ought we to do so as to decide whether something is true? We should have to inquire whether it is *true* that an idea [thought?] and a reality, say, correspond in the specified respect. And then we should be confronted by a question of the same kind, and the game could begin again. So the attempted explanation of truth as correspondence breaks down, and any other attempt to define truth also breaks down. For in a definition certain characteristics would have to be specified. And in application to any particular case the question would always arise whether it were *true* that the characteristics were present. So we should be going round in a circle.<sup>36</sup>

---

<sup>36</sup> Scott Soames, in his class at the CUNY Graduate Center, Fall 1996, noted (and presented in a prepared manuscript for his forthcoming book) that the way the argument is set up, it might look as if the claim is that any analysis is circular, since, if truth is analyzable (e.g., as T), then, in inquiring whether or not p is true, we must also inquire whether p is T, and, because of the putative transparency of truth, whether p is T is true, etc. As Soames notes, this argument is not sound: if an analysis is circular, it follows that establishing that something falls under the analysans will involve establishing that it also falls under the analysandum; however the converse implication does not hold.

I think that this is an interesting argument; I think that it brings out an important point; but I do not think that, in the end, it poses much of a worry. For ease of explication, I rephrase it as follows.<sup>37</sup> We are to consider the following series:

p	Tp	TTp	TTTp	...
p	Φp	ΦΦp	ΦΦΦp	...

- (i) In the first series, each member has the same content as any other; this is the transparency property
- (ii) If  $\Phi$  is an analysis of truth, then it must share the transparency property
- (iii) Either  $\Phi$  is a genuine predicate, containing concepts of its own, or it is a manufactured synonym of T.
- (iv) If the latter, then no analysis is offered.
- (v) If the former, then it cannot share the transparency property, and, again, by (2), is no analysis of truth.
- (vi) So there is no possibility of analysis of truth.

Before we discuss the argument, we should ask after Frege's positive conclusion. In other words, given his skeptical conclusion vis a vis the possibility of analyzing truth, what positive claim does he make? It is not easy to tell from this text, for, after establishing to his satisfaction that truth is "*sui generis* and indefinable," he asks, rhetorically, (1988, p. 37) "May we not be dealing here with something which cannot be called a property in the ordinary sense at all?" This leaves it open whether *truth* is a primitive and indefinable property, or whether it is no property at all and truth ascriptions are redundant.

Hence, we might conclude that Frege's argument, if it goes through, establishes that we cannot analyze truth, opening the door to a primitivist view *or* to a redundancy view. Given the problems we have found with the redundancy view, we might take

---

<sup>37</sup> This very nice way of setting it up comes from Blackburn, 1984, p. 228. Blackburn offers no solution; he merely suggests that Frege's own view might fall victim to the problem. I will not comment on his claim; however, I should point out that if my solution is correct, Frege has to worry no more than does anyone else, which is not much.

Frege's argument as supporting a primitivist line, one according to which truth is a simple, indefinable property, possessed primarily by propositions, and only derivatively by utterances, beliefs and sentences. This is the line adopted, more or less, by Horwich (1990, 1995, 1997), who claims that truth is a simple indefinable, "logical" property whose "nature" is exhausted by the collection of all non-paradoxical instances of (DS), and that it is, therefore, susceptible of no deep analysis. Hence, it might be that Frege's argument vindicates a minimalist position; indeed, Horwich's property of truth might capture the sense in which truth is "something which cannot be called a property in the ordinary sense at all."<sup>38</sup>

It is of historical interest, what route Frege took, but it is not of much interest what route this argument should lead him to take, for, as far as I can tell, the crucial premise, (v), rests on a mistake. Let's rephrase (v) as (v'),

(v') If  $\Phi$  is a genuine predicate, containing concepts of its own, then it cannot share the transparency property, and, again, since, if  $\Phi$  is the analysan in an analysis of truth, it must share the transparency property, it cannot be the analysan in an analysis of truth.

The problem with (v') is that the second conjunct seems to be false. That is, it seems possible for  $\Phi$  to play a large role in an analysis of truth, i.e., for truth to be analyzed as  $\Phi$ , though  $\Phi$  does not share the transparency property. More to the point, suppose that truth is analyzed as  $\Phi$ . Such would be a *theoretical* identification of the property of truth, much like (27) involves a theoretical identification of the property of gold.

However, the

---

<sup>38</sup> It might also be that Frege's argument supports a substantivist primitivist view of the sort once championed (albeit briefly) by Russell and Moore.

(27) Gold = the element with atomic number 79

issue of transparency regards the *concept*, not the *property*: though gold *is* the element with atomic number 79, one who understands the content of a gold ascription—e.g., “that is gold”—does not have to understand what is expressed by “that is the element with atomic number 79.”

If the Transparency thesis is correct, then transparency is a property of the concept of truth, for the claim that ‘is true’ is transparent is compatible with, and actually motivates, the redundancy theory, according to which there is no property of truth. Whether or not truth enjoys an analysis is a fact about the property of truth. Hence, if truth is  $\Phi$  then  $\Phi$  has whatever properties the property of truth has (and vice versa), but it does not follow from this that the concept of truth and the concept of  $\Phi$  must be identical, and, hence, it does not follow that, even if truth is  $\Phi$ , and the concept of truth is transparent, the concept of  $\Phi$  is transparent.

The argument fails, but it forces us to recognize an important distinction: the distinction, made famous by Kripke (1972/1980), between the concept of X and the property X. The former is what we grasp when we grasp what is expressed by an “X”-containing sentence; the latter is what is invoked if X plays an important theoretical or explanatory role. As we will see, this distinction is among the most important that we can make, when we are considering the notion of truth.

#### 1.4 The Quinian Theory<sup>39</sup>

---

<sup>39</sup> I should note that in speaking about the Quinian theory I am not intending to capture the truth about disquotationalism, as that has been developed by Field and Leeds. Those are among the primary views that I will consider; in this section, I am trying to provide the historical underpinnings.

According to Ramsey, the purpose of the truth predicate is to allow us to formulate assertions in various styles. Quine sees the truth predicate as playing a more definite role. Quine picks up on what we might think of as Frege's second use of the truth predicate, and has developed it into what is now considered the fundamental or basic function of that predicate. According to Quine, the truth predicate allows us to talk about reality by talking about language; it is useful as a device of *semantic ascent*. Furthermore, semantic ascent is useful insofar as it allows us to formulate generalizations that we could not formulate otherwise, unless our language came equipped with substitutional quantification or some equivalent device, which it does not. Here is Quine (1970, pp. 11-12)

We can generalize on 'Tom is mortal', 'Dick is mortal, and so on, without talking of truth or of sentences; we can say 'All men are mortal'... When on the other hand we want to generalize on 'Tom is mortal or Tom is not mortal', 'Snow is white or snow is not white', and so on, we ascend to talk of truth and of sentences, saying 'Every sentence of the form "p or not p" is true...

If we want to affirm some infinite list of sentences that we can demarcate only by talking about sentences, then the truth predicate has its use. We need it to restore the effect of objective reference when for the sake of some generalization we have resorted to semantic ascent.

Quine's discussion of the function of the truth predicate, that is, of the conceptual role of "is true", serves as a sort of fulcrum for the deflationist positions of Leeds (1978, 1995),

Williams (1986), Horwich (1990, 1995, 1997), and Field (1994), who take Quine to have captured completely the *raison d'être* of truth.

What can we say beyond this? We cannot look to Quine's claim about the utility of the truth predicate, in order to determine what sort of account of truth he countenances, for his claim about the utility of the truth predicate is compatible with a wide range of positions on the nature of truth. So what does Quine think? For one thing, he thinks that utterances are the proper truth bearers,<sup>40</sup> but in saying this, he is not committing himself to the claim that a true utterance has some property not possessed by a false one. He notes,

there is some underlying validity to the correspondence theory of truth, as Tarski has taught us... To ascribe truth to the sentence ['snow is white'] is to ascribe whiteness to snow; such is the correspondence in this example. Ascriptions of truth just cancel the quotation marks. Truth is disquotation. (1990, p. 80)

He continues, anticipating the obvious objection

So the truth predicate is superfluous when ascribed to a given sentence; you could just utter the sentence. But it is needed for sentences that are not given... the truth predicate is an intermediary between words and the world. What is true is the sentence, but its truth consists in the world's being as the sentence says... The disquotational account of truth does not define the truth predicate... for definition in the strict sense tells how to eliminate the defined expression for every desired context in favor of previously established notation. But in a looser sense the

---

<sup>40</sup> Cf. Quine, 1990, p. 77.

disquotational account does define truth. it tells us what it is for any sentence to be true... we understand what it is for the sentence 'Snow is white' to be true just as clearly as we understand what it is for snow to be white... 'True' is transparent. (Ibid., pp. 81-82)

We recognize, in Quine's claim, a repeat of the Fregean line: the disquotational account<sup>41</sup> does define truth in some sense, and, hence, 'true', on that sort of use, is transparent; but a provision of every instance of the disquotational schema does not define 'true' in the sense that to define X is to provide a formula that will allow us to eliminate it from all contexts, for, among other things, we have the blind ascriptions that give truth its point.<sup>42</sup>

Quine notes that the traditional philosophical question about truth, namely, What is truth?, is usually understood as asking for something more than the disquotational account. Nevertheless, Quine thinks that the disquotational account is completely satisfactory: first, it makes no false predictions—there is no disputing that, to take Quine's example, 'snow is white' is true if and only if snow is white; and, second, it is complete—it explicates clearly the truth (falsity) of every clear and unambiguous sentence. Of course it does not provide anything deeper—for example, a unified account of the truth of every sentence of the language—but, to use an expression from Quine (1970, p. 46), it provides all that can be asked for: clear intelligibility without (because of the paradoxes, blind uses, etc.) full eliminability.

---

<sup>41</sup> Just for the record, according to Quine, the truth predicate *disquotes* a sentence S if the form '\_\_\_ is true if and only if \_\_\_' comes out true when S is named on the left-hand side and written on the right.

<sup>42</sup> Furthermore, the disquotational schema does not hold for all substituends—the natural language truth predicates (e.g., 'true') are only partially disquotational.

It has been thought by some (e.g., Price, 1988 and Horwich, 1990) that Quine endorses a redundancy account of truth, and the quote provided above is cited as evidence to that effect. It has also been thought that his endorsement of Tarski's account suggests that he accepts that view of truth. It is not clear that he accepts a redundancy account—he never mentions Ramsey or redundancy in the relevant passage—and he compares his view with Tarski's,<sup>43</sup> which suggests that those, too, should be distinguished.

In fact, it seems that the best way to understand Quine's view is as a sort of deflationary correspondence account. He seems to hold<sup>44</sup> that it is possible to maintain the truth-as-correspondence view, so long as we do not make any attempt to provide a *reductive* account of the notion of correspondence. What this amounts to is less than clear, but it is clear what value Quine sees in Tarski's work. According to Quine, the real virtue of Tarski's work is that it charts the internal structures of given languages (or linguistic systems), provides an effective method for correlating each sentence of an object language with a sentence from a metalanguage that is taken to state its truth condition, and, thereby, sheds light on linguistic structure, which is important to such endeavors as translation (if possible) and interpretation. However, Quine holds, the concepts that Tarski employs are no more clear than the concept of truth.

Field (1972) held that instances of schema T adequately clarified the meaning of "true" (for ordinary, non-philosophical purposes), but he rejected instances of the T schema as acceptable philosophical explications of that notion, because they did not explain that notion in terms of other, more primitive, physicalistically acceptable words-

---

<sup>43</sup> Cf., Quine, 1990.

world connections. Field applauded Tarski's work, but thought that he did not complete the physicalist project: in order to complete the project, Tarski has to provide an empirical account of primitive denotation (or reference/satisfaction), i.e., to render it physicalistically acceptable.

Quine disagrees, for he does not see the reduction of truth to primitive denotation (and structure) as providing a more illuminating account of the nature of truth; in fact, he suggests (1960, 1969) that Tarski's procedures, insofar as they explicate truth for a language relative to an interpretation, actually presuppose, rather than illuminate, the concept of truth. Hence, Quine holds that Field was wrong to think that, unless we got a physicalistically acceptable account of primitive denotation, the notion of truth would have to be abandoned, for he questions the extent to which a reductive account of truth (for a language) makes that notion any clearer than it already was (for that language).

It is not of too much independent interest to us exactly where Quine fits. We will call his view *Quinian disquotationalism*, and will take the positions developed by Field (1994a,b) and Leeds (1978, 1995) to be variants of it. What does matter, however, is that Quine claims that Tarski's account of truth, i.e., in terms of satisfaction, together with his inductive definition of satisfaction, render both accounts equally virtuous, because Tarski's account "explains satisfaction of each specific sentence, but it does not provide a translation of 'x satisfies y' with variable 'y'." (Quine, 1990, p. 86) Hence, like the disquotational account, though it explains truth for every (closed) sentence of a given language, it does not explain it in application to variable 'y', which is to say that it does not provide some one general theoretical identification or definition of truth, i.e., a

---

<sup>44</sup> See Quine 1968, 1970, 1990.

synonym, in simple English, that can everywhere replace ‘true’ or ‘satisfies’ in ‘x is true’ or ‘x satisfies y’.

What this means is that, although Tarski showed us how to provide an (materially adequate) explicit definition of truth, using only non-semantic concepts that could already be expressed in the language for which truth was being defined, syntactic concepts applying to expressions of the language, and elementary set theory (including logic), he did not provide a general *analysis* of truth,<sup>45</sup> i.e., a general account of what it is for a sentence to be true. This is one point of intersection between the Disquotational theory and the Tarskian theory: neither provides the “deep” account of truth expected by (e.g.) the traditional correspondence theorist. Furthermore, like the Disquotational theory, the Tarskian theory makes no false predictions; and, again, like the Disquotational theory, it, too, is (taken to be) as complete as can be expected. For these reasons, as far as Quine is concerned, it makes no difference which one adopts; the results will be the same either way.

How can this be? How can Quine adopt such an attitude? Although I am not sure that anyone has troubled to answer this *prima facie* troubling claim, it is important, for, among other things, it sheds much needed light on Quine’s rejection of Field’s claims about the incompleteness of Tarski’s account. The answer is simple, but it is worth noting, nevertheless: both the Quinian and the Tarskian accounts aspire to nothing more than extensional correctness, and there can be more than one extensionally correct account of a notion. Since both satisfy the desiderata that concerned Quine, each was

---

<sup>45</sup> Indeed, as Soames (in discussion) has pointed out, it is plausible to claim that he didn’t provide any analysis of truth at all. As Soames sees it, Tarski provided a stipulative definition of a new technical term,

just as acceptable as is the other. Furthermore, with extensional correctness as the goal, Quine could take Field's criticisms of Tarski to be misplaced. If all we are going for is an extensionally correct definition, then Field is wrong to claim that a reductive analysis of primitive denotation is needed.

Let's consider Quine's claim that there was no real basis for choice between his view and Tarski's. I think that there is a sense in which Quine is right and that there is a sense in which he is wrong. He is clearly right that, given his two criteria, together with the fact that the goal was extensional correctness, there is no real basis for choice. However, this just shows that his criteria are not discriminating enough—there are significant differences between a Tarskian and a Disquotational account of truth, differences that might lead a deflationist to accept one over the other. In order to see this, we must have, on the table, some semblance of a Tarskian account, only then will we be able to compare the two.

### 1.5 A Tarskian Deflationary Theory<sup>46</sup>

We will let Tarski speak for himself:

We begin with a simple problem. Consider a sentence in English whose meaning does not raise any doubts, say the sentence 'snow is white'. For brevity, we denote this sentence by 'S', so that 'S' becomes the name of the sentence. We ask ourselves the question: What do we mean by saying that S is true or false?

The answer to this question is simple: in the spirit of Aristotelian explanation, by

---

'true', a restricted predicate that is designed to do all the theoretical work that is needed over a limited part of the language.

<sup>46</sup> It is crucial to note that not all philosophers understand Tarski as providing a deflationary account of truth. Field (1972), and Devitt (1996) both view Tarski as providing an incomplete correspondence theory.

saying that S is true we mean simply that snow is white, and by saying that S is false, we mean that snow is not white. By eliminating the symbol 'S' we arrive at the following formulations:

(1) 'Snow is white' is true if and only if snow is white

(1') 'Snow is white' is false if and only if snow is not white.

Thus (1) and (1') provide satisfactory explanations of the meaning of the terms 'true' and 'false' when these terms are referring to the sentence 'snow is white'. We can regard (1) and (1') as partial definitions of the terms 'true' and 'false', in fact, as definitions of these terms with respect to a particular sentence. Notice that (1), as well as (1'), has the form prescribed by the rules of logic, namely the form of logical equivalence. It consists of two parts, the left and the right side of the equivalence, combined with the connective 'if and only if'.

(1969, p. 64)

Tarski's project was to show how we could explicitly define truth predicates for fragments of various formal languages. Hence, like Quine, and unlike Frege, Tarski took the primary truth bearers to be sentences, or symbols. Furthermore, Tarski's aim was not to define an inter-linguistic or transcendent concept of truth, but to provide truth definitions for (fragments of) specific languages—*true-in-L*, rather than *true*.

Tarski accomplished two tasks. First, he showed how we could define in first-order logic a notion equivalent in function to purely disquotational truth, given that we had sufficient ontological resources.<sup>47</sup> Second, he showed how to define truth-in-a-

---

<sup>47</sup> This is where the Quinian questions come in. As Field notes (1986, p. 64), what Tarski actually did was "almost the same thing: show how to define [in purely first-order logic] substitutional quantification with

model, a non-disquotational, relativized, purely formal notion, one that is applicable to languages that we do not understand.

It is this second task, the task of providing an account of truth-in-a-model that has become important in the characterization of the difference between inflationist and deflationist accounts, and, as is important for present purposes, in the characterization of the difference between a Tarskian inflationist and a Tarskian deflationist account.<sup>48</sup> In order to show this, we need a little terminology. We assume, with Tarski, that all of the relevant languages that we might consider can be formulated as first-order languages. The interpretation of a language, then, is to be understood as the specification of a domain, denotations for names and extensions for atomic predicates. Associated with every interpretation,  $I$ , is a relation-in-extension,  $T$ , the truth relation under  $I$ , as we will call it. We introduce ‘R’ to pick out the “standard” interpretation of English, i.e., the one whose domain contains everything, and which assigns Bill Clinton to ‘Bill Clinton’, the set of presidents to ‘president’, etc. We will call the associated truth relation ‘T’. R and T are the standard interpretation and truth relation for English, respectively.

According to Leeds (1978), the divide between Tarskian deflationists and their inflationist (or, as we are calling them, *substantivist*) cousins (i.e., inflationist Tarskians) comes down to the answer to this question: what makes R and T the standard interpretation and truth relation for English? In other words, What makes the standard interpretation and truth relation relevant and some other(s) irrelevant? The substantivist explains what makes R and T relevant by pointing to facts about how we use the

---

formulas as the substitution class.” Both disquotational truth and what Tarski provided could serve as a device of infinite conjunction; hence, modulo the obvious restrictions, they are functionally equivalent.

language; hence, the substantivist accepts a *use-dependent* notion of truth conditions: what makes it the case that the standard interpretation is *correct* regards facts about how we use the language. If we had used the language differently, one of the non-standard interpretations would have been correct. One important consequence of this is that, although the standard interpretation is correct, this is a contingent fact.

By contrast, the deflationist holds that there is no explanation for why it is that R and T are the standard interpretation and truth relation for the language; R and T are simply stipulated to count as relevant. This means that there is no fact about how we use the language that explains the correctness of R and T, from which it follows that R and T, though in some sense correct, are not correct in virtue of some fact about our language use. Since R and T are not correct in virtue of any facts about us or about how we use the language, it follows that sentences do not have their truth conditions in virtue of the way in which we use the language, which means that, following Ayer, the deflationist offers a *use-independent* account of truth conditions. As we will see in chapter 4, this has interesting consequences for the plausibility of a deflationist account. For now, we note only that, when we restrict our attention to the Tarskians, what distinguishes inflationist from deflationist Tarskians is how they explain what makes a given assignment of truth conditions correct.

Let's return to Tarski. Tarski showed how, using only non-semantic concepts already expressible in the language for which truth was being defined; syntactic concepts that applied to expressions of that language; and set theory (including elementary logic), we could define a truth predicate for each language. By providing limited truth

---

<sup>48</sup> In the characterization of the difference, I follow Leeds (1995). In so doing, of course, I do not intend

definitions for specific languages, Tarski thought that he was capturing our ordinary concept of truth. At this point, such a claim might sound ridiculous: how could Tarski's limited truth predicates, which only apply relatively, and then only to fragments of the languages over which they are defined, do that?

In order to answer this question, we have to get a bit more clear on how Tarski conceived his project. First, using the limited material mentioned above, Tarski provided *materially adequate* truth definitions for the relevant languages (or language fragments), which is to say that he provided truth definitions that entailed Schema T, for

$$(T) s \text{ is true in } L \leftrightarrow p$$

each sentence of the fragments of the languages over which the truth predicate was defined, where 's' is replaced with a quote name of a sentence, 'p' is replaced by a sentence of the metalanguage that means the same as the named sentence, and '↔' is understood as the material biconditional.

Second, Tarski (1944) clearly thought that it was a fact about our ordinary concept of truth that the claim that a sentence is true is materially equivalent to the sentence itself. Put another way, per Tarski,<sup>49</sup> our ordinary concept of truth accords trivial status to (M-T), and it is because of this that we can be confident that any

$$(M-T) s \text{ means in } L \text{ that } p \rightarrow s \text{ is true in } L \leftrightarrow p$$

definition of truth that entails schema (T) applies to all and only the truths of our language.

---

to signal agreement with his conclusions.

<sup>49</sup> The answer comes from a number of places, most notably, Soames (lecture notes). However, it is important to note that Horwich has taken (M-T), or something like it, very seriously; he has claimed that it captures the heart of intentionality. We will return to this issue in chapter 4.

One still might not be satisfied. After all, one might have thought, the job is not merely to define a predicate that applies to all and only the truths of (a fragment of) our language, i.e., to define a predicate that is coextensive with ‘true’ over a fragment of the language. Rather, the goal is to provide an *analysis* of the concept of truth, to say what it is for a sentence to be true, and one might wonder in what sense Tarski has done this.

This is a legitimate worry, and answering it will pave the way to a better understanding of contemporary deflationist positions, especially those of Field and Leeds. The first thing to note is that Tarski has, indeed, said what it is for a sentence to be true (in a language)—he has said what it is for ‘snow is white’ to be true (in English), for example: ‘snow is white’ is true in English if and only if snow is white. This, of course, is not what a critic means when she launches the aforementioned worry, however. What she means is that Tarski has not provided a simple, well-formed ordinary English equivalent to *s is true*, for variable *s*, that can everywhere replace *s is true*. This is certainly correct, but it is irrelevant: that is not what Tarski set out to do. For one thing, if we take him at his word (1944, p. 342), he seems to think that Aristotle has already done that. His interest is to improve on what Aristotle has provide—as he says (Ibid.), to “look for a more precise expression of our intuitions.”

How ought we to understand Tarski’s project? The best way to understand in what sense a Tarskian account of the concept of truth is an analysis of that notion is to see Tarski as providing a *stipulative definition* of a restricted truth predicate, one that is designed to do the theoretical work that needs to be done. Such an analysis is adequate if the newly introduced truth predicate does everything that we can legitimately expect a

truth predicate to do, and if it does everything that we can legitimately expect a truth predicate to do, it is not clear what more we could ask of it.

What this means is that, in order to determine whether or not the newly introduced truth predicate is up to the job, we should determine what the job is, i.e., what explanatory work we expect out of our notion of truth—only then can we see if Tarski's truth predicate is capable of doing the work that we can expect it to do. Is this correct? Must the Tarskian deflationist do this? I do not think so. Surveying the explanatory role of the truth predicate is necessary only if we do not have reason to believe that the newly introduced truth predicate fails to capture the *meaning* of 'true-in-L', for our language, L. Considering the quote that began this section, it is obvious that that is precisely what Tarski aimed to do: he thought that each instance of schema T partially defined truth and that, therefore, a definition that entailed all of these instances would capture what we mean by 'true', as that is applied to our own language.

In short, if we want to show that an analysis is adequate, we must show that it yields a notion that can do all of the theoretical work we expected of the notion analyzed. One way to do this is to provide the job description for the unanalyzed notion and, painstakingly, to show that the new notion does everything we expected, or had a right to expect, of the old notion. This is a dangerous strategy, as it is always open to the critic to complain that we have no way of knowing that the list is complete. A better strategy, then, is to show that our new notion captures the *meaning* of the old one, i.e., that we have a notion that captures our concept of truth. In this case, we do not need to provide a job description; if we are right, we are ensured that the new notion will be adequate. This, I believe, is what Tarski (and, as we will see, Horwich) recognized.

Has Tarski succeeded?<sup>50</sup> Does the Tarskian truth predicate capture the meaning of our ordinary truth predicate? We have already heard from Quine, who was somewhat skeptical about the claim that Tarski's methods renders the ordinary notion of truth any more clear than it is made if we restrict ourselves to the disquotational account. We turn to Scott Soames, for, according to him (1995, pp. 252-255), there is reason to worry.<sup>51</sup> Although I do not think that Soames is right, it will be useful to go through a version of his argument. Recall Schema T, and consider its Tarskian cousin, Schema TT, where 'T' contains the definiens of Tarski's formal definition of truth.

(Schema T)  $s$  is true in  $L \leftrightarrow p$

(Schema TT)  $s$  is T in  $L \leftrightarrow p$

Consider Abe, an ordinary speaker of  $L$ , which happens to be English, and suppose that  $s$  is a sentence of English that Abe does not understand, for example, (28).<sup>52</sup> Suppose that Abe is told that (28) is true in English.

(28) The punctilious pugilist got pummeled.

According to Soames, Abe can learn something about the meaning of (28), just in virtue of the fact that he understands our ordinary truth predicate. For example, from his understanding of the truth predicate, according to Soames, he knows (29).

(29) 'The punctilious pugilist got pummeled' is true in English  $\leftrightarrow$  The punctilious pugilist got pummeled.

---

<sup>50</sup> There are numerous ways to read Tarski, varying in degree of inflation. My counter argument is intended to be in defense of a deflated Tarskian, which means, in particular, a Tarskian who does not see truth as playing a vital role in semantics or the theory of meaning. If we can discuss no such Tarskian, then there is no sense in which there is a Tarskian Deflationist position. Since I (and others, e.g., Field) think that there is a Tarskian Deflationist position (of at least some sort), I (and others) think that we can discuss such a Tarskian deflationist position. Is this Tarskian Tarski? It is not clear (although I tend to doubt it). More importantly, it is not important, at least not for present purposes.

Furthermore, he can use his knowledge of (29) to infer that ‘the punctilious pugilist got pummeled’ does not mean in English (30). However, the same does not hold for T;

(30) The punctilious pugilist did not get pummeled.

if Abe does not already understand (28), by being given the relevant instance of Schema TT, he is not going to be able to infer that that ‘the punctilious pugilist got pummeled’ does not mean in English (30), for, though (29) is trivial and a priori, (31) is not, and,

(31) ‘The punctilious pugilist got pummeled’ is T in English  $\leftrightarrow$  The punctilious pugilist got pummeled.

therefore, is not available for drawing conclusions about meaning.

As Soames admits, he assumes that the truth predicate plays a substantial role in the theory of meaning, and that the notion of truth is required in semantics, as well as in the theory of meaning. The problem is that a notion of truth that did what Soames expects would not be a deflationary notion of truth, at least not as we are understanding Tarskian deflationism. This is an important point, and one that we will return to later, for Soames clearly believes that a deflationary account of truth is compatible with a substantive account of meaning, even one in which the concept of truth plays an important role.<sup>53</sup>

Soames notes an intensional connection between meaning and truth, which connection is reflected in the alleged triviality, a priority and necessity of (M-T), which triviality, a priority and necessity is said to flow from the conceptual connection that exists between the notions of meaning and truth. He assumes, then, that “statements

---

<sup>51</sup> We will return to this issue later, when we discuss the conceptual role of truth.

<sup>52</sup> This is not Soames’ example, so he bears no responsibility for its silliness.

about truth conditions do tell us something about meaning” (1995, p. 253), and that “if the notion of truth really is required in semantics and the theory of meaning, then Tarski cannot be regarded as having provided an adequate analysis of it.” (1997, p 24) The problem, as I see it, is that if truth does have an explanatory role in semantics or the theory of meaning, then, not only a Tarskian deflationist theory but, any thorough-going deflationist theory would be inadequate.

More to the point, if sentences get their truth conditions as a matter of stipulation, then the fact that a sentence *S* is true iff *p* is wholly independent of how we use *S*, and, in particular, cannot yield any facts about what we mean by it. More to the point, if *R* and *T* are merely stipulated to be relevant, then the fact that *S* has the truth conditions that it has cannot tell us anything about *S*'s meaning, for *S*'s truth condition is not connected with its meaning in the way that Soames imagines: *S* does not have the meaning that it has in virtue of having the truth condition that it has. I imagine that a substantivist will balk at this view, but it is very clear that it is a consequence of Tarskian deflationism.

We can go further. According to Soames, it is the job of a theory of meaning to explain how it comes to pass that the sentences of our language have the truth conditions that they have. It does this in two stages: at stage one, the theory of meaning assigns propositions to sentences; at stage two, it assigns truth conditions to propositions, and, derivatively, to the sentences that express them. However, as should be obvious, when combined with Tarskian deflationism, such a view is patently false, and I assume that the reason why is obvious: if there is no explanation for how it comes to pass that *S* gets the

---

<sup>53</sup> As we will see, Soames is an advocate of what I have called propositional deflationism, the view that is metaphysically, though not metasemantically, deflationist.

truth condition that it has, then it is certainly not the job of a theory of meaning to explain such facts.

Since Soames identifies a sentence's meaning with what it does with its expressing the proposition that it expresses (excepting cases of indexicality), it follows that a sentence has the truth conditions that it has in virtue of its having the meaning that it has, which is to say that we can explain, in meaning-theoretic terms, what makes it the case that, for each (relevant) non-indexical sentence, *S*, of the language, *L*, there is a proposition, *P*, such that *S* is true in *L* iff *P* is true. Generalizing, then, and, again, relying on the theory of meaning, we can say what it is for one truth definition to be correct, and another one incorrect. If this is the case, then we are right to wonder in what sense Tarski's account, understood in the way that Soames would have us understand it, is deflationist. Soames does not tell us, but it is easy to see that if we expect our notion of truth to yield substantive conclusions about the meanings of sentences, we are not conceiving of truth in a suitably deflationary way.<sup>54</sup>

We have rejected Soames' putative problem as not being a problem for the Tarskian deflationist. However, we should note that if he is right about the conceptual role of truth—i.e., about the conceptual connection between meaning and truth—then, if the conditions of adequacy on a theory of truth include that (e.g.) the Tarskian deflationist capture all of our pre-theoretic intuitions about truth, we would be able to draw another conclusion from Soames' argument—to wit, that the Tarskian cannot be a deflationist. Of course, as previously noted, the conditions of adequacy that Tarski accepted do not include that we capture the pre-theoretic notion of truth (N.B., if any

such thing exists), so we need not pursue this issue any further.<sup>55</sup> The Tarskian deflationist is offering a new notion that is supposed to do the theoretical work we have a right to expect of the old. It remains to be seen, the Tarskian deflationist will argue, whether there is any theoretical work that needs to be done that cannot be done if the intensional connection between meaning and truth is severed.

Are there remaining problems with the Tarskian account? One potential problem is that, if the disquotationalist is right, it might be more substantial than the pre-theoretical notion of truth that it aims to replace. For example, on the Tarskian approach, compositional principles are built into the inductive characterization of truth, which is to say that, on the Tarskian approach, to give a theory of truth, one must adduce compositional principles, and, if our ordinary notion of truth is tied to the disquotational schema, it is not clear that it requires that, or that it is even possible to adduce such an account.<sup>56</sup>

We can put this in a slightly different way. Tarski (1933, p. 155) considered and rejected the following definition of truth: “a true sentence is one which says that the state of affairs is so and so, and the state of affairs indeed is so and so.” Why he rejects it is of some interest: for one thing, he does not think that it is too vague or unclear, as he does other peoples’ definitions. In fact, he says that it is this definition that he is trying to capture. There are, however, two reasons he rejected it. First, it has the semantic ‘says’

---

<sup>54</sup> As the reader has certainly guessed, by ‘a suitably deflationary way’ I mean a *thorough-going deflationary way*.

<sup>55</sup> Of course one might claim that Tarski had an inadequate set of conditions of adequacy, and that may very well be true. However, establishing that would involve raising different issues from the ones that we are considering here.

<sup>56</sup> As Field, 1994, p. 269 notes, a Tarskian truth theory must be compositional, whether it provides an explicit definition of truth, an inductive definition of truth, or an axiomatic account.

in the second conjunct, and, hence, it is not acceptable by the lights of physicalism. Second, it does not show how grammatical structure affects truth value. Hence, it was because of his physicalist and model-theoretic interests that he rejected the aforementioned account, which is to say that Tarski's commitments generate from sources that the deflationist might reject, and, therefore, might be more substantive than the pre-theoretical notion that it aims to replace.

Ought the deflationist to be a Tarskian deflationist? I think that, as the deflationist's see it, the answer is that she should not be a Tarskian deflationist. There are two reasons for this. The primary reason for thinking this is that, as the deflationist sees it, there is no compelling reason for thinking that Tarski's model-theoretic ambition need be the deflationist ambition, for there is no reason to think that a satisfactory notion of truth will include the principles of composition, and other model-theoretic machinery, that Tarski employs. More generally, in the interest of keeping her explanatory machinery to a minimum, the deflationist will reject the employment of the machinery necessary to run Tarski's recursive definition of truth.

The second reason follows from a criticism, which has been leveled by people from Max Black (1948, pp. 60-61), to Quine (1960, p. 24), to Putnam (1981, pp. 128-129), to Hartry Field (1986, p. 71), and it is sometimes called *the vacuity objection*. Field states this criticism in what I think is its least contentious form, when he notes that Tarski's theory is compatible with every actual and possible theory not easily refuted by counter-example. I am not sure if this is true, for, if the points mentioned above are correct—namely, if the deflationist, and, in particular, the disquotationalist, holds that compositionality fails in some areas of discourse—then it might not be true that Tarski's

theory is as easy-going as Field says, and, what is far more important for our purposes, it might not be the case that it is compatible with other forms of deflationism (from which I do not think Field would infer that those other forms, most notably, disquotationalism, are easily refuted by counterexample).

There are other objections of note, such as Blackburn's (1984, p. 266-267) relativity objection, i.e., that Tarskian truth is not a translanguistic notion, and, therefore, is no more revealing about the nature of truth than are the following series of concepts of the concept of proper legal verdict: proper-legal-verdict on Monday, proper-legal-verdict on Tuesday,... I believe that we have said enough to make a quick reply to this criticism. It is not a criticism of Tarski's theory, so much as it is a criticism of Tarski's goals. Hence, in order properly to assess the strength of Blackburn's criticism, a criticism that can equally well be leveled against any deflationist account that takes truth to be an *immanent* notion, we must know, in more detail than we have provided thus far, what it is that we should expect a theory of truth to do.

### **1.6 Comparison of a Tarskian and a Disquotational View**

We have seen most of the ways in which a Disquotational view can be contrasted with a Tarskian deflationist view, but I will highlight a couple that I take to be of central importance. First, a Tarskian truth predicate is defined for languages that we don't understand, whereas disquotational truth (at least in its most pristine form) is not. On the pure disquotational view, one can understand truth ascriptions only insofar as one can understand the utterances to which truth is ascribed. Another way of putting this, less mysteriously, I believe, is that the purely disquotational notion of truth is only applicable

to utterances that are meaningful to the truth ascriber, and then, for me to say that an utterance is true is for me to say that, as I, the truth ascriber, understand the utterance, i.e., as I interpret it, it is true—as Field (1994a,b) puts it, it is true-as-I-understand-it.

Second, the advocate of the purely disquotational notion thinks that the only reason we have a truth predicate is to serve our logical needs. This is not to deny that when we say that an utterance is true, we are not ascribing some property to it,<sup>57</sup> rather, it is to say that the *raison d'être* of the truth predicate is to serve our logical needs. The Tarskian truth predicate has a more limited area of application. For example, we cannot apply it to some (or all) sentences containing semantic terms, and, more worrisome, we cannot apply it to sentences containing unrestricted quantifiers. The result seems to be the disquotational truth predicate is better suited to serve our logical needs—the needs for which we have a truth predicate, according to the deflationist—than is the Tarskian truth predicate.

Furthermore, as previously noted, the Tarskian explains the truth conditions of complex sentences on the basis of those of atomic sentences, together with certain rules of composition. Hence, compositional principles are built into the Tarskian's inductive characterization of truth (Cf. Soames, 1984 and Field 1972, for very lucid expositions of Tarski's formal project). Here is a place where, as previously mentioned, the Tarskian departs from the disquotational schemata, for it is not clear that it will follow from the generalized version of the disquotational schema, (T\*),

---

<sup>57</sup> I address the question of whether or not the disquotationalist holds that there is a property of truth in the next chapter. As I show, there is good reason to hold that she does not. More on this later, however.

(T\*)  $(\forall p)$  (“p” is true iff p)<sup>58</sup>

that compositional principles will hold in all areas of discourse. This is not to say that they will not hold in all areas, of course; however, for the disquotationalist, the question of whether or not compositional principles will hold in a given domain can be answered only by appeal to the inferential roles of the concepts that populate that domain, together with the disquotational function of ‘true’. What this means is that, as the disquotationalist sees it, compositional principles are a negotiable commodity: we employ them where we need them, but there is no a priori guarantee that they will be applicable to all parts of the language. Hence, whatever one’s stance on the importance of compositional principles, i.e., whether or not one thinks them indispensable in our explanation of language learning (Cf. Davidson), or in our explanation of the productivity and systematicity of language (and thought) (Cf. Fodor (1987, 1991) for the *pro* view and Schiffer (1987, 1993) for the *con* view), it seems clear that they do not furnish an *a priori* requirement for the acceptability of an account of truth.

This last point needs to be pressed, for I think that it is under-appreciated: the difference between a Tarskian and a Disquotationalist view of the principles of compositionality come down to this: for the Tarskian it follows from the way in which truth is characterized that the truth conditions for complex expressions are arrived at compositionally. The Disquotationalist might end up with an equivalent view, i.e., she might end up employing a rich set of compositional principles as well, but the difference between the two is that, in the case of the disquotationalist, if she does employ the same

---

<sup>58</sup> Note that here the quantifier is understood substitutionally. As Field (1994) notes, one could forego the substitutional quantifier and opt, instead, for schematic variables, in which case the quotes should be taken to be corner quotes.

rich set of compositional principles that the Tarskian employs, then, restricting our attention to the concept of truth, it is an accident, for, as she makes clear, it is not a condition on the adequacy of a characterization of truth that this be the case.

### 1.7 Concluding Remarks

The purpose of this chapter is to set out what must be set out if we are properly to evaluate deflationism. I cite seven worries that might drive a philosopher to adopt deflationism, all of which derive, in one way or another, from the discussion in this chapter.

- worries about the *nature* of truth—in particular, that it has no nature.
- worries about the *property* of truth—that there is no such thing, or that, though there is a property of truth, it is a special, non-substantive, no-naturalistically reducible property.
- worries about the nature of the *concept* of truth—that it is thin, i.e., that truth is a mere logical notion.
- worries about the *possession conditions* for the concept of truth—that they are exhausted by our disposition to accept instances of one or another version of the Disquotational Schema—“p” is true if and only if p..
- worries about the *conceptual role* of the truth predicate—that a grasp of the concept of truth is not presupposed by a grasp of other of our central semantic, doxastic, epistemological or metaphysical concepts
- worries about the *explanatory role* of truth—that it plays no role in any of our theories of the world; in psychological explanation; in epistemology; or in logic.

- worries about truth (and reference) conditions—that we can provide no systematic connection between the meaning of a sentence (expression) and its having the truth (reference) condition that it has.

These “worries” constitute the various claims made by the contemporary deflationists.<sup>59</sup>

It is to that class of positions that we turn next.

## **Chapter 2: Contemporary Deflationism**

Traditional accounts of truth aim to give an elucidation of truth, where, by an *elucidation*, I mean an attempt to exhibit constitutive non-analytic connections between truth and other properties,  $\Psi$ ; i.e., connections expressible as

For all  $x$ ,  $x$  is  $\Phi$  iff (or: if; or: only if)  $x$  is  $\Psi$ ,

that is, a definition, analysis, or explication of truth that tells us what truth is or what it should be like if it is to function as it must in our logical, empirical or mathematical theories. Of course this way of setting things up presupposes that the property of truth has a significant theoretical role to play, or, at the very least, that there’s something useful, illuminating or non-trivial that we can say about truth, but this is a claim uniformly accepted by those who are in the business of providing substantive analyses and explications of truth.

As we have seen, Frege denied that an elucidation of truth can be given. This was one of the things that led him to adopt a redundancy view of truth. Our present concern is not with historical versions of deflationism, however; our present concern is with

---

<sup>59</sup> I should note that the contemporary deflationists that I am concerned with include, among others, Hartry Field, Paul Horwich, Stephen Leeds, W.v.O. Quine, Huw Price, Simon Blackburn, Vann McGee, and, in particular, do not include Robert Brandom, whose work, though very interesting, has made it into this study in only a few places.

contemporary deflationism, the view held, in one form or another, by Field, Leeds, Horwich, McGee, and many others. What does the contemporary deflationist say about truth? In particular, since, as we will see, unlike the substantivist, the contemporary deflationist denies that we can explicitly define truth, we have good reason for wondering what, according to the deflationists, truth is.

There are numerous contemporary deflationary accounts of truth, and there are important differences among them. Nevertheless, there seem to be four theses that are held by nearly all deflationists, and on which any discussion of contemporary deflationist views of truth should focus. These are:

(T1) The entire description of truth is captured by the instances of simple schemata such as “*p*” is true iff *p* and it is true that *p* iff *p*, etc. Depending on what the deflationist is willing to quantify over, she will accept some or all of (S1).

(S1)	Truth for sentences	(ES) “p” is true iff p
	Truth for propositions	(EP) The proposition that p (= <p>) is true iff p
	Truth for beliefs	(EB) A belief that P is true iff p
	Truth for utterances	(EU) An utterance of “p” is true iff p

What is important about the schemata listed above is that they cannot be transformed into closed formula expressing candidate elucidations or analyses.<sup>60</sup>

The deflationists also assent to (T2) - (T4)

(T2) What is expressed by sentences of the form ‘It is true that p’, or ‘p is true’, where ‘p’ is a stand-in for a proposition, belief, sentence or utterance, is, in some sense, trivially equivalent to p, which is to say that they are necessary, a priori consequences of each other.

---

<sup>60</sup> We will say more below about what exactly the deflationist takes the instances of the disquotational schemata to amount to. Doing so will provide us with insight into the way in which the deflationist thinks about the nature of truth.

- (T3) What is expressed by a biconditional of the form ‘p is true iff p’ is necessary, a priori and capable of being known, on the basis of linguistic knowledge, by anyone who understands it.<sup>61</sup>
- (T4) The explanatory role of truth is exhausted by the function of the concept of truth, which is to serve as a device of generalization

(T1) - (T4) regard central features of our thinking about truth: (T1) regards the deflationist stance towards the shape *theories* of truth should take; (T2) and (T3) regard the cognitive significance of the truth predicate; and (T4) regards the function of truth, i.e., why we have a predicate that does what “true” (and cognates) does. Though (T1) - (T4) are answers to centrally important questions about truth, even a cursory glance at these theses should yield a rather surprising result. If the deflationist aims to replace the substantivist accounts of truth with her more austere, favored account, we would expect her account to yield an answer to the traditional question, (Q), that is either more

(Q) What is truth?

intuitively appealing or, at least, better suited to the purposes for which we have a notion of truth. However, (T1) - (T4) do not yield an answer to (Q).

After unpacking the theses, (T1 - (T4), we will turn to (Q), and we will consider some of the most popular contemporary deflationist theories, to see how they answer it. In the course of our discussion, I will develop some arguments which, if successful, provide us with reason for favoring a more liberal version of deflationism. This is the

---

<sup>61</sup> On one way of understanding the notions employed in (T4), we might say that a proposition p is knowable a priori iff there is a way of grasping p such that, if a person grasps p in that way, she knows it without appeal to empirical evidence. Furthermore, for propositions p,q, we might say that q is an a priori consequence of p iff for any mode of access to p, A, there is a mode of access to q, A', such that it is possible, in principle, to reason deductively from p under A to q under A' without appeal to empirical evidence. Compare Field's claim that we accept an instance of (DS) in virtue of the cognitive equivalence between the left- and right-hand sides, i.e., his claim that we accept instances of (DS) because our inferential procedures license a (fairly direct) inference from the left- and right-hand sides.

version that I will assume in the ensuing chapters, though I intend my arguments to hold more generally, extending to the versions of deflationism that I find less compelling.

## 2.1 Regarding (T1)

The deflationary theories of truth are to be contrasted with inflationary, or substantivist, theories of truth, where those are normally understood as theories that attempt to provide definitions of truth in terms of notions like correspondence, coherence, facts, representational fit and the like. In fact, as many see it,<sup>62</sup> the focus of the debate over the nature of truth is between those philosophers who believe truth to be a causal-explanatory notion, and, therefore, a notion defined in terms of correspondence or representational fit, and those who deny this claim. Of course the deflationist does not think that truth is a causal-explanatory notion that can be defined in terms of correspondence or representational fit, but a substantivist can deny this claim as well, so I think we should seek a different characterization of the debate, i.e., one that is more inclusive.

In fact, I think that attempts to distinguish deflationists from substantivists on the basis of the claim that truth can be defined in terms of one or the other of the aforementioned substantive notions is not broad enough to distinguish deflationist theories of truth from substantivist theories of truth, generally. Let's call this conflict—the one between deflationists and substantivists who believe that truth can be defined in terms of substantial notions like correspondence, coherence, representational fit, etc.—

---

<sup>62</sup> Stephen Schiffer suggested to me that the real debate was between the deflationist and the philosopher who accepts that truth is a causal-explanatory notion. Although I agree that this is an important debate and that for many philosophers, such as Hartry Field, these are the only two (plausible) options, I do not think

the conflict between *deflationists* and *inflationary substantivists*. The deflationist disagrees with the inflationary substantivists, of course, but she disagrees with others as well,<sup>63</sup> which means that there is another dimension to the debate between deflationists and substantivists. This facet of the debate does not regard the *content* of a definition of truth, i.e., the nature of the definiens, but the *form* of a definition of truth—in particular, whether or not we can or should provide an *explicit* definition of truth. I take this to be a central issue in the debate over the nature of truth, though it is not stressed in much of the literature on the issue. We turn to this contrast next.

### 2.1.1 Deflationary Theories vs. Tarskian Theories<sup>64</sup>

In order to bring out this contrast, it will be useful to distinguish deflationary theories of truth from Tarskian theories of truth.<sup>65</sup> We can distinguish the two by reference to the role played in the respective theories by instances of the disquotational schema, (DS)

(DS) “p” is true iff p.

---

that this characterization of the debate is inclusive enough to capture some of the important differences between deflationist theories of truth and substantivist theories of truth.

<sup>63</sup> If all substantivists adopt some form of inflationary substantivism, we might say that deflationists disagree with inflationary substantivists both on the claim that we can define truth in terms of some non-semantic notions like those mentioned above *and* on other claims that do not, or do not directly, regard the definiens of a definition of truth.

<sup>64</sup> Didn't we already do that? No. We distinguished the Tarskian view from the Disquotational view. If what we say in section 2.1.1 is correct, we have another significant distinction between a Tarskian and a Disquotational view of truth.

<sup>65</sup> I suspect that, due to certain moves that have been made by deflationists—e.g., restricting the domain of the truth predicate to a finite base—what I will say may not be true of all deflationist positions. I do not think that that matters, because I think that what I say here captures the heart of deflationism, even the heart of deflationist theories that do not accept it.

According to the Tarskian, instances of the disquotational schema<sup>66</sup> provide a *minimal condition of adequacy* on a definition (or, theory) of truth, which definition is given in a *metalanguage*, the language *in* which truth is being defined, to be contrasted with an *object language*, which is the language *for* which truth is being defined. That is, as the Tarskian sees it, a theory of truth, to be adequate, must entail all instances of (DS).<sup>67</sup> By contrast, the deflationists do not take instances of (DS) to provide a minimal condition on the acceptability of a theory of truth; instead, they take the instances of (DS) to *be*, or to *constitute*, a theory of truth.

The Tarskian attempts to provide an explicit definition of truth in the metalanguage, whereas the deflationist does not. Instead, as Horwich (1990, 1998) describes it, the deflationist attempts to provide an *implicit definition* of truth in the object language. On her view, *truth* is implicitly defined by all instances of (DS), which instances are taken to be ordinary, object-level sentences. This points to an important difference between the two approaches, a difference that seems to make deflationism attractive to many people. By Tarski's theorem on the undefinability of truth, it is not possible (or, at least, not likely) to define truth in the language for which truth is being defined.<sup>68</sup> It is for this reason that the so-called (e.g., by Soames, 1997, although he does

---

<sup>66</sup> For our purposes, the distinction between the disquotational schema and the T-schema will be irrelevant. The difference, for what it is worth, relies on the need for the use of quote names in the former, though not in the latter.

<sup>67</sup> Of course it should not entail an instances of (DS) for every sentence that can be formulated in the object language, for there are *paradoxical* sentences that could be formulated in the object language. There are various strategies for restricting the range of the metalinguistic truth predicate, but we will not go through those here. It is sufficient to note that no Tarskian would hold that all instances of (DS) formulatable in the object language need be, or should be, entailed by a theory of truth.

<sup>68</sup> Incidentally, this raises an interesting question about inflationary substantivist theories of truth. It appears that those theories attempt to define truth within the languages to which truth is to apply. How can they do this, given Tarski's theorem on the undefinability of truth? One way would be for them to renounce Tarski's condition of adequacy. This issue should be investigated more fully, for it might provide a disincentive for the development of inflationary substantivist theories of truth. Arnie Koslow suggested that

not discuss this issue) “Semantic theories of truth”—the theories of Belnap and Gupta, Kripke, and Tarski, to name the most famous—cannot be described in (the theory of) the object language. In order to describe them, you need recourse to a *metatheory* that is “stronger” than the theory of the object language.<sup>69</sup> This means that if we want an explicit definition of truth, we will need to give it for a language that is weaker than our own.

If we do not go in for an explicit definition of truth, i.e., if we can be content with an implicit definition of truth, then we do not need recourse to the meta-theoretical resources that the inflationist needs—Tarski employed second-order quantification—which means that we can obtain a notion of truth for our own language, rather than for one that is weaker than our own.<sup>70,71</sup> The deflationist avails herself of nothing more than the truth predicate and the governing principles for it or its use,<sup>72</sup> and, in particular, she does not employ the special apparatus necessary explicitly to define truth. This is one feature of deflationism that is thought to make it especially attractive: the deflationist employs only what is needed, and nothing more than is needed, to implicitly define (or,

---

the substantivists have simply neglected questions raised by Tarski’s theorem. If this is correct, and my survey of the literature has provided me with reason for believing that it is, then we might uncover a compelling reason for rejecting substantive (e.g., causal) theories of truth.

<sup>69</sup> As Richard Heck, Jr. (1997, p. 545, n. 20) points out, Tarski is sloppy about the distinction between languages and theories formulated in those languages. Much of what we say about languages—that they are inconsistent; that they imply things—holds, in fact, only of theories formulated within those languages. In what follows, I, too, will not be too careful about respecting this distinction, except where it matters, of course.

<sup>70</sup> I would like to thank Arnie Koslow for pointing this out to me, and for having innumerable discussions that have led me to appreciate this important feature of deflationist theories.

<sup>71</sup> Of course this means that the deflationist will have to come up with some way of dealing with the paradoxes, but I leave that issue to the side.

<sup>72</sup> Here is where deflationist positions tend to veer off. Field, for example, has what might be called an *inference-rule deflationism*. Rather than taking as his account all instances of (DS), he takes as his account principles such as *from s infer “s” is true* and *from “s” is true infer s*. Of course he needs something more than this—for example, the deductive theorem or conditional proof—but I take it that he can employ that as

as we shall describe it, to *axiomatize*) a notion of truth that serves our practical and theoretical needs.

### 2.1.2 The Status of the Instances of (DS) and Semanticalism

So, one dimension along which we might distinguish deflationist views of truth from their substantivist counterparts regards the theories of truth that they endorse. Whereas the substantivist retreats to a metalanguage, in which she attempts to explicitly define a notion of truth, the deflationist makes do with the resources provided in the object language, i.e., she tries to define truth in the language for which the notion is being defined. As a result (e.g., when the language is a natural language such as English), she is unable to provide an explicit definition of truth, which means that one significant difference between deflationist and substantivist theories regards the sorts of definitions of truth that they attempt to provide.<sup>73</sup> This raises an interesting and important question, one that has not been discussed by many philosophers and that must be addressed if we are to pull out of the deflationists an answer to (Q): what is the status of the instances of (DS)? The way in which this question has been taken, it is usually asked in reference to (T2) or (T3), where it is taken to involve the cognitive significance of ‘true’. As should be obvious, the question of the cognitive significance of the truth predicate is not what we are after here. We want to know what kind of a theory of truth the deflationist is providing. We have it that the instances of (DS) formulate the deflationist’s theory of

---

well. What results from these principles, however, is all instances of (DS) (limited by certain other restrictions that are irrelevant to our present concerns).

<sup>73</sup> Note that the Tarskian also provides an implicit definition of truth, for she provides an inductive definition of truth. It is important to notice that there are very serious differences between the inductive definitions of truth provided by the Tarskian and the implicit definitions of truth provided by the deflationist. Using familiar methods developed by Tarski, the Tarskian can convert her inductive definition of truth into

truth,<sup>74</sup> and that, in particular, they are not considered by the deflationist to be conditions of adequacy on a theory of truth, so how are we to take the instances of (DS)? As previously mentioned, Horwich, (*passim*, 1990, 1998) tells us that the instances of (DS) *implicitly define* truth, but what does that mean?

As I understand the proposal, the claim that the instances of (DS) implicitly define truth amounts to the claim that they *axiomatize* truth, i.e., that each instance of (DS) plays the role of an *axiom* in the theory of truth. These axioms do not define truth in the usual way, and, contrary to what some people seem to think, they do not even fix the extension of the truth predicate.<sup>75</sup> So what do they do, and, given what they do, what does the deflationist think truth is? As noted, the deflationist takes the instances of (DS) to axiomatize truth, or, as we might say, to *characterize* our notion of truth, though the collection of axioms taken together do not provide us with the resources necessary to explicitly define truth. After all, they are just ordinary sentences of the object language. What results from the deflationist's theory, then, is an undefined, primitive notion that is *axiomatized* by all instances of (DS).<sup>76</sup> Truth, for our deflationist, is a primitive,

---

an explicit definition of truth. The deflationist cannot do so. Hence, although some might say that both the deflationist and the Tarskian provide implicit definitions of truth, this is liable to obfuscate the issue.

<sup>74</sup> Excepting certain deflationists, of course (see previous footnote).

<sup>75</sup> The claim that the totality of T sentences does not fix the extension of truth was noted by Tarski, 1935, p. 258). As has been noticed, failure to note this fact has led people to think that the deflationist—our deflationist, at any rate—has access to the set of all instances of (DS), which is not the case.

<sup>76</sup> Incidentally, for those, like Field, who do not overtly adopt this view, the status of the instances of (DS) is really no different: Field's inference rules characterize truth and the instances of (DS) are "more or less "analytic" or "logically true"..., by virtue of the cognitive equivalence of the left and right sides." (1994b, p. 405). We should note that Field's use of "cognitive equivalence" is inappropriate. By 'cognitive equivalence' Field means nothing more than that our inferential rules license fairly directly inference from either one to the other. As Field is well aware, most people take cognitive equivalence of "s" and "s' is true" to be the basis for the inferences from the one to the other, rather than taking it to consist in it. Hence, Field's claim that instance of (DS) are logically true by virtue of their cognitive equivalence is a bit misleading. At any rate, the claim that the instances of (DS) are logically or conceptually true because we can infer the one side from the other is not too much different from our deflationist's claim that they are logically or conceptually true in virtue of the fact that they follow from the meaning of the truth predicate.

undefined, notion that is axiomatized by the instances of (DS). It is because it is primitive and undefined that no further metatheory need be provided in our characterization of it.<sup>77</sup>

We have our deflationist's answer to (Q):<sup>78</sup> truth is a primitive, undefined notion. This raises a very interesting, and, given the recent history of deflationism, somewhat ironic issue that I can only touch on here. The claim that truth is a primitive, undefined notion is dangerously close to a view that Field (1972) aimed to resist—*semanticalism*. Semanticalism, as that was characterized by Field (1972, p. 358), is the view that

semantic phenomena (such as the fact the “Schnee” refers to snow) must be accepted as primitive *in precisely the way that electromagnetic phenomena are accepted as primitive* (by those who accept Maxwell's equations and reject the ether; and in precisely the way the biological phenomena and mental phenomena are accepted as primitive by vitalists and Cartesians.)” (italics added)

The worry about semanticalism, dualism, etc., although generated largely by physicalist leanings, can be appreciated even from the perspective of someone not ultimately wedded to physicalism (e.g., someone who is not confident that there is a coherent idea behind the physicalist's claims). It is the worry that we will have to claim that truth is a primitive explanatory notion, meaning that, like spin, charm and

---

Again, then, I do not think that there are relevant differences between the form of deflationism that I am sketching and, e.g., Field's version of the view, though, of course, there are differences (such as, for example, the truth bearers).

<sup>77</sup> I would like to thank Volker Halbach for very interesting e-mail discussions of this point.

<sup>78</sup> As we will see, we have *one* deflationist answer to (Q). The other alternative, which is just a variant of this one, will be discussed shortly.

charge, it will end up on Fodor's fabled list of ultimate and irreducible properties.<sup>79</sup> This is something that no good physicalist could accept, but it is not something that anyone else should be happy with either, at least not without a substantial argument. So, I will take the consensus to be that, absent any pro-semanticist arguments, semanticism is something that everyone—deflationist and inflationist alike—might want to avoid.

Leeds (1978, p. 120) claimed that we do not have to provide a physicalistically acceptable reduction of semantic terms in order to avoid semanticism. Although I think that Leeds' argument rests on a confusion,<sup>80</sup> and, so, although I do not think that he is correct for the reasons that he gives, I think that, nevertheless, he is probably right. Here's why. If our deflationist takes truth to be an axiomatized notion, then she is committed to the claim that if there really is a property of truth then it is a primitive, indefinable property. Hence, *prima facie*, at least one half of the semanticist's claim must be accepted by the deflationist

That notwithstanding, the deflationist is not committed to semanticism.

Semanticism, as described by Field, is the view that truth is accepted as a primitive notion *in precisely the way that electromagnetic phenomena are accepted as*

---

<sup>79</sup> Cf. Fodor, 1987, p. 97.

<sup>80</sup> Here's what Leeds says: "the Tarski sentences, which axiomatize the notion of truth, are by no means a complicated or recondite axiomatization: the possibility of moving from this axiomatization to the explicit definition was always latent in the logical structure of our language, though it took a Tarski to discover it." Now, as is well known, Tarski discovered that we could not define truth within L for L, so, strictly speaking, Leeds is wrong. Given the Tarski sentences, *qua* object-level sentences, and assuming that there are an infinite number of them, you cannot get an explicit definition of truth; you need to move to a metalanguage in order to get an explicit definition of truth for any natural language.

What is the price of Leeds' error? Well, if deflationism is understood as we are understanding it, then the deflationist cannot get an explicit definition of truth. If, as we have claimed, she must make do with an axiomatization (or something equivalent to that), then it is fair to say that, to the extent that she is willing to countenance a property of truth, her notion of truth is a primitive property, which means that, in one sense, semanticism might not be avoided after all.

*primitive*. Although Field does not tell us how electromagnetic phenomena are accepted as primitive, it is not too hard to guess: electromagnetic phenomena are *both primitive and explanatory*. What Leeds should not have argued is that the deflationist avoids semanticalism in virtue of the fact that, appearances to the contrary notwithstanding, she can explicitly define truth, for this is not the case. What he should have argued was that the deflationist avoids semanticalism because, even if she is committed to there being a primitive and indefinable notion of truth, she is not committed to their being a primitive, indefinable notion of truth that plays a role in any of our theories or explanatory practices, and, thereby, requires us to say, in other words, what truth is.

Indeed, as noted in (T4), according to the deflationist, the explanatory role of truth is exhausted by the function of the truth predicate—‘is true’ (and cognates). If the playing of this role does not require us to provide an explicit definition of truth, as the deflationist believes that it does not, then the deflationist will avoid semanticalism, though she accepts the ontological half of that doctrine. Assuming the deflationist can establish that truth fills no other explanatory roles (an issue to which I will turn in chapter 3), she can rest easy, for, though she admits that truth is a primitive, undefined property (N.B., if she countenances a property of truth at all), she denies that it is explanatory, and, therefore, she is not committed to semanticalism, as that was characterized by Field.

The sense in which the deflationist can really avoid semanticalism might be obscured by the fact that she seems committed to the ontological half of the doctrine, the half that might appear worrisome. In other words, if she accepts that there is an

irreducible property of truth, then, even if she denies that it plays an explanatory role akin to the role played in physical theory by electromagnetic phenomena, isn't she still committed to a vicious semanticalism? A parallel: would substance dualism be made acceptable by the rider that our non-physical Cartesian souls are explanatorily inert?

I think that this criticism misses the point, and I think that it is important to see why. Although the deflationist might countenance a primitive property of truth, following Horwich,<sup>81</sup> she might adopt a minimalist view of properties, or at least, of the property of truth, as in (MP).

(MP) A property, *F*, is minimal iff, for objects *a* and *b*, it is possible that both *a* and *b* exemplify property *F*, though there is no intuitive common explanation for why both *a* exemplifies *F* and *b* exemplifies *F*.<sup>82</sup>

If she does adopt a minimalist view of the property of truth, she will avoid semanticalism, for, though she might accept that, in some sense, there is a property of truth, she will not countenance the sort of property that *could* play a substantive explanatory role in either our theories or our practices.<sup>83</sup>

So, our deflationist takes truth to be a primitive, undefined notion that is characterized by the axioms of the theory of truth, which axioms are instances of (DS), i.e., ordinary sentences of the language for which truth is being defined.

---

<sup>81</sup> Horwich (1990, 1998) claims that truth is a *minimal* or *thin* property, but he does not say exactly what he takes minimal properties to amount to. In what follows, I attempt a characterization with which I believe that he would be happy.

<sup>82</sup> Of course the deflationist will adopt a minimalist view of property exemplification (such a view is entailed by what I call propositional deflationism) according to which *a* exemplifies *F* iff *a* is *F*. Hence, we can say that *a* exemplifies *F because a* is *F* and *b* exemplifies *F because b* is *F*, but that there is no common explanation for what makes it the case that *a* is *F* and what makes it the case that *b* is *F*.

Of course a theory based on (DS), unless restrictions are imposed, will be inconsistent. To simplify matters, Horwich (1990, 1998) has proposed to exclude troublesome sentences. McGee (1992) has shown that there is really no good way to do that, i.e., that it is not clear that there is a method that Horwich can employ to single out a unique maximal consistent set of set of sentence. So, to simplify matters, following Soames (1998), we'll take our deflationist's theory of truth to consist of all instances of (DS) that do not include either the truth predicate or any cognate notion.

## 2.2 Regarding (T2) & (T3)

According to (T2), what is expressed by sentences of the form 'It is true that p', (T2) What is expressed by sentences of the form 'It is true that p', or 'p is true', where 'p' is a stand-in for a proposition, belief, sentence or utterance, is, in some sense, trivially equivalent to p, which is to say that they are necessary, a prior consequences of each other.

or 'p is true', where 'p' is a stand-in for a proposition, belief, sentence or utterance, is, in some sense, trivially equivalent to p, which is to say that they are necessary, a prior consequences of each other.<sup>84</sup> Something like (T2) explains why we accept instances of the disquotational schema, according to Field (1994a,b) and Soames (1998), each of whom declare their versions of (T2) partially to capture the heart of a deflationist view.

We should note that the claim that "p" and "'p' is true" are trivially equivalent (in this sense) is different from the claim made by the redundancy theorists that they are

---

<sup>83</sup> This modal claim will be important, when we distinguish deflationism from Moorean or Russellian primitivism.

<sup>84</sup> See footnote 59 for a characterization of this notion. As previously mentioned, Field's characterization would be a bit different, though I cannot see that the difference is of much consequence for our purposes.

*cognitively equivalent*, i.e., in the sense that they have the same meaning or express the same proposition, which is why I shy away from the terminology that Field (1994a,b) favors.<sup>85</sup> As mentioned previously, evidence that they are not cognitively equivalent (in the usual sense of that notion, to the extent that there is one) is easy to find: to take one obvious example, a person can understand the one without understanding the other (if, for example, she does not yet “have” the concept of truth). However, it seems that, as our notion of *trivial equivalence* is understood, assuming that the deflationist can successfully make it out, the claim that the left- and right-hand sides of the instances of (DS) are trivially equivalent should be relatively uncontroversial.

We turn, next, to (T3), where, as previously noted, per Soames (1997, p. 2), (T3) What is expressed by a biconditional of the form ‘p is true iff p’ is necessary, a priori and capable of being known, on the basis of linguistic knowledge, by anyone who understands it.

a proposition p is knowable a priori iff there is a way of grasping p such that, if a person grasps p in that way, she knows it without appeal to empirical evidence. According to Field and Soames, (T3) follows from (T2). Horwich (1997, 1998) independently accepts

---

<sup>85</sup> I also shy away from Field’s terminology because I find it to be somewhat misleading. Field’s claim that instances of (DS) hold because of the cognitive equivalence of the left- and right-hand sides should not be confused with the Redundancy theorist’s claim that this is the case. The redundancy theorist might argue as follows:

- (a) “p” and “p’ is true” are cognitively equivalent (in some sense that I will not specify); so, given (b),
- (b) If two propositions, p and q, are cognitively equivalent then our inferential procedures license a fairly direct inference from the one to the other,
- we have it that (c)
- (c) Our inferential procedures license a fairly direct inference from the right-hand side of an instance of (DS) to the left hand side of that instance

and this explains our acceptance of the instances of (DS). By contrast, Field takes the cognitive equivalence of p and q to consist in facts like (c), and, on that basis, he claims that the cognitive equivalence of “p” and “p’ is true” explain our acceptance of the instances of (DS). Hence, his claim that the cognitive equivalence of the left- and right-hand sides of the instances of (DS) explains why we accept them is a bit misleading.

(T3), claiming that instances of the disquotational schemata are *conceptually fundamental*, by which he means “that we accept [them] in the absence of supporting argument: more specifically, without deriving them from reductive premises...” (1997, p. 95-96)

We should note that Horwich takes this to be “the fundamental thesis of deflationism,” (1997, p. 96) a thesis that he contrasts with the *inflationist* thesis, according to which our acceptance of the instance of (DS) depends on our belief in the obtaining of a certain representation relation holding between (e.g.) the sentence, “snow is white,” and the fact that it represents iff it is true, to wit, that snow is white.<sup>86</sup> Although we cannot quibble with Horwich’s characterization of (T3) as a fundamental principle of his favored version of deflationism, it is probably worth pointing out that it does not mark any great divide between inflationist and deflationist positions. Horwich’s claim that instances of (DS) are *conceptually basic* is consistent with any reasonable theory of truth that the substantive inflationist would put forward. The reason is that the substantive inflationist is putting forward a theory of the *property* of truth, not, or not exclusively, a theory of the *concept* of truth. The inflationary substantivist who claims that truth consists in correspondence with the facts, coherence with one’s other beliefs, or whatever, can accept whatever facts about our concept of truth Horwich presents, for the claim that instance of (DS) are *conceptually* fundamental is compatible with the claim that they are not *explanatorily* fundamental, where, as I understand it, the claim that the

---

<sup>86</sup> Of course Horwich means to contrast his view with all substantivist views, and not just with representationalist views. Hence, more generally, we might take him to contrast his view with any view that takes it that our acceptance of an instance of the disquotational schema depends on our acceptance of an instance of some other (reductive) premise stating a relation between the left- and right-hand sides of the

instances of (DS) are explanatorily fundamental is the claim that we cannot explain the fact that they hold by reference to any more basic facts, such as might be adverted to in an inflationary substantivist's theory of truth. Since, as we will discuss, the instances of (DS) can be conceptually basic without being explanatorily fundamental, insofar as the substantivist is concerned with the property of truth, and, hence, with the question of whether or not instances of (DS) are explanatorily fundamental, her position is compatible with the claim that they are conceptually basic.

This raises two important and interconnected issues. First, how could it be that the instances of (DS) are conceptually basic though they are not explanatorily fundamental? Second, given our answer to that question, we might wonder why, if instances of (DS) really can be explanatorily non-fundamental while being conceptually basic, we should care about (T3), or, more specifically, about the claim that they are conceptually basic? In other words, why is (T3) an important deflationist thesis? Actually, I think that (T3) is a very important thesis, though I do not think that it has received the sort of attention that it deserves, as I hope to show.

How can instances of (DS) be explanatorily non-fundamental and conceptually basic at the same time? The answer turns on the explanatory role of truth. If truth is conceptually basic in the way Horwich has claimed, then, if the inflationist is to support her contention that it is explanatorily non-fundamental, she must carve out an explanatory role for truth to play, and, in fact, she must carve out an explanatory role that calls for a more-than-deflationary notion of truth.<sup>87</sup> The reason should be obvious.

---

disquotational schema. As we will see, this way of characterizing things is confused, and, indeed, it obfuscates the real issue between deflationists and their inflationist detractors.

<sup>87</sup> Cf. my discussion of Gibbons' argument, in 3.4.3.

Suppose that instances of (DS) are conceptually basic. Suppose, as well, that whatever explanatory role truth plays, its playing that role is compatible with the deflationary account of truth. In that case, there would be no reason to seek a more-than-deflationist notion of that property, no reason to expect a more-than-deflationary account of the property of truth.

Given that, we have an answer to our second question. If the deflationist is right about (T3), and, in particular, if Horwich is right that instances of (DS) are conceptually basic, then if the substantivist is to support her position, she must adduce an explanatory role for truth to play, for, assuming that instances of (DS) are conceptually basic, it is only if truth plays an explanatory role that requires our characterization of that notion to go beyond the axiomatization provided by the instances of (DS) that the substantivist has reason to expect to be in the offing a more-than-deflationist account of the property of truth. Hence, though (T3) is not incompatible with an inflationist view, it establishes an important theoretical burden for the inflationist, the burden of adducing an explanatory role for truth (or truth conditions) that cannot be played by a deflationary notion.

### **2.3 Regarding (T4)**

Finally, we turn to (T4).

(T4) The explanatory role of truth is exhausted by the function of the concept of truth, which is to serve as a device of generalization.

So, (T4) combines two claims, (T4a) and (T4b).

(T4a) The truth predicate is a device of generalization.

(T4b) The explanatory role of truth is exhausted by the function of the truth predicate.

(T4a) comes from Quine (1970, p. 12), who writes:

By calling the sentence [“snow is white”] true, we call snow white. The truth predicate is a device of disquotation. We may affirm the single sentence by just uttering it, unaided by quotation or by the truth predicate; but if we want to affirm some infinite lot of sentences that we can demarcate only by talking about the sentences, then the truth predicate has its use. We need it to restore the effects of objective reference when for the sake of some generalization we have resorted to semantic ascent.<sup>88</sup>

To illustrate the truth predicate’s logical function, suppose that, for whatever reason, you develop the belief that Allison is an infallible guide to everything. How do we capture the content of your belief? It isn’t enough to say that you believe everything that she says, for your belief is more radical than that. If we had time and access enough, we could express your belief by running through an infinite conjunction beginning like this: if Allison says that cats are dangerous then cats are dangerous & if she says that rainbows are dangerous then rainbows are dangerous, ad infinitum. If we’re willing to quantify over propositions, we can capture this infinite conjunction as a universal quantification: for every proposition  $x$ , if what Allison says =  $x$  then  $x$  is true. This, in turn, can be put more informally as Everything Allison says is true. It is in this sense that the concept of truth is useful as a device of generalization.

(T4b) goes beyond (T4a), and it distinguishes the deflationist’s view of the explanatory role of truth from the view of the explanatory role of truth that has been adopted by some substantivists. According to (T4b), the explanatory role of truth is

exhausted by the function of the truth predicate. Hence, (T4b) entails that once you have described the function of the concept of truth, you have said all that there is to say about the explanatory role of truth. Given (T4a), (T4b) entails that the explanatory role of truth is exhausted by the logical role of the truth predicate.

As might be obvious, Given (T2) and (T3), and given the distinction that we have drawn between the claim that something is conceptually basic and the claim that it is explanatorily fundamental, it should be clear how important (T4) is to the deflationist's project—(T2) and (T3), when combined with (T4), plus the distinction between being conceptually basic and being explanatorily fundamental, support, though they do not entail, (T1), or some variant thereof. It is for this reason that I take (T1), or something like it, to capture the heart of contemporary deflationism. Hence, in what follows, we will assume a version of deflationism that includes (T1), though in general the issues that I raise will be insensitive to the precise formulation of deflationism that we adopt.<sup>89</sup>

#### 2.4 Disquotationalism vs. Deflationism Revisited

We noted previously that (T1) - (T4) did not contain an answer (Q), the What is truth? question. However, we noted that we could develop an answer to (Q) out of (T1). It is important to note that there are other deflationist answers to (Q). In what follows, I

---

<sup>88</sup> Cf. our discussion of Frege's distinction between the two contexts of use for a precursor to the Quinian view.

<sup>89</sup> So, for example, the issues will hold for Field's inference-rule version of deflationism, for Horwich's infinite propositional theory and for Sosa's (1993) finitary propositional theory, i.e., the theory that includes propositions like

- For all sentences, *s*, EXP(*s*) is materially equivalent to TRUE(*s*), where EXP is a function from non-propositional truth bearers—for example, utterances—to the propositions they express, and TRUE is a function from a putative truth bearer to the proposition to the effect that they are true.
- For all propositions, *p*, *p* is materially equivalent to TRUE (*p*)
- For all beliefs, *b*, if *b* is a belief that *p* then *p* is materially equivalent to TRUE(*b*).
- For all utterances *u*, if *u* is an utterance of *s*, then EXP(*s*) is materially equivalent to TRUE(*u*)

will discuss disquotationalism and some of its variants, and I will try to show why a deflationist might be inclined towards a more liberal, less austere view. One of the issues that I will raise will involve directly (Q), so we will see at least one other answer to that question that can be given by the deflationist.

Impressed by the logical function of the truth predicate, deflationist such as Quine (1970, pp. 11-12) have taken truth to be a *mere* device of abbreviation, denying that there's any property of truth at all.<sup>90</sup> We will call this view *disquotationalism*. There are reasons for thinking that disquotationalism is problematic, but I will not catalogue those here.<sup>91</sup> Instead, I want to point to two reasons for thinking disquotationalism is not the most attractive version of deflationism. First, however, I will run through the reason for claiming that the disquotationalist denies that there is a property of truth, for this claim often goes unrecognized.<sup>92</sup> The reason is simple: If the truth predicate is *merely* an abbreviatory device, then if 'vrai', 'verdad' and 'true' express properties (and we will assume that they do), they express *different* properties, in which case none expresses *the* property of truth. Of course we might try providing a functional account, letting *true*<sup>93</sup> pick out the property of having the aforementioned abbreviatory function, but in that case, it would be a property of truth predicates, rather than a property expressed by truth predicates. Hence, if truth is immanent, there is no property of truth.

---

together with the appropriate logic and rules of inference.

<sup>90</sup> This is not a claim that is often made by the disquotationalists, though we will show that, in fact, it is entailed by their view.

<sup>91</sup> For an extensive and very good discussion of some of the problems with disquotationalism, see David (1994).

<sup>92</sup> The claim that there is no property of truth is not the same as the point made by Field to the effect that a disquotational notion of truth is *immanent*, rather than *transcendent*, although it is clearly related to that point. I also ignore general claims to the effect that there is no property of truth, that might be made by the nominalists, who claim that there are no properties at all.

This by itself is not a problem for the disquotationalist, of course, though it might strike some as counter-intuitive.

The real problem with claiming truth to be a mere device of abbreviation, as Gupta (1993) has pointed out, is that claims of abbreviation entail claims of meaning equivalence, and we have good reason for thinking that generalizations such as (32)

(32) Everything Gupta says is true

are not equivalent in meaning to the corresponding generalization, (33)

(33) If Gupta said that disquotationalism is doomed then disquotationalism is doomed,

and if he said that revisionist semantics is the way to go, then revisionist semantics is the way to go, ... ad infinitum

In fact, they are neither conceptually nor logically equivalent to the infinite conjunctions that they are said to *express*. Very briefly, they are not conceptually equivalent, as Soames (1997,1998) has noted, because it is possible to understand truth generalizations without grasping the correlated infinite conjunction; and they are not logically equivalent because universal generalizations are not logical consequences of the collection of all of their instances.

This raises an interesting question: if uses of the truth predicate in a type two context (e.g., (32)) are neither logically nor conceptually equivalent to the infinite conjunctions and disjunctions that they are said to express, in what sense can they be said to express such infinite conjunctions and disjunctions? As I understand it, one answer to this question is that they “express” infinite conjunctions and disjunctions in the sense that they play the same role in reasoning as would be played by infinite conjunctions and

---

<sup>93</sup> Just to remind the reader, I use boldface to indicate that what is being used is an expression for a

disjunctions. What allows them to play this role is that, though such generalizations are neither logically nor conceptually equivalent to the infinite conjunctions and disjunctions they are said to express, they are equivalent with respect to their 'true'-free consequences, i.e., with respect to the statements, sentences, or what have you that they imply that do not contain the truth predicate. So, consider our sentence, (32), and the sentence (34), where 'p' is restricted to 'true'-free sentences.

(34) If Gupta said "p" then p.

First, restricting ourselves to 'true'-free sentences, (32) implies all sentences (34), though the collection of sentences (34) does not imply (32). However, the collection of sentences, (34), does implies itself. Hence, (32) and the collection of all sentences, (34), are equivalent in their 'true'-free consequences.

Nevertheless, I think that a thorough-going deflationists should opt for a version of deflationism that recognizes a property of truth, and, anyway, it is this sort of deflationist that I will be primarily concerned with in what follows, though, as before, I suspect that what I say will hold for other deflationists, including, among others, the disquotationalists. Our deflationist, then, is what we will call a *property deflationist*, a deflationist who accepts that there is a property of truth and accepts (T1) - (T4).

The property deflationist recognizes a transcendent, rather than an immanent, notion of truth (Cf. Quine, 1986, p.19, for the distinction), and, hence, she is not eliminativist about truth. She accepts that sentences like (32) express universal propositions.<sup>94</sup> As previously mentioned, the property deflationist holds that there is a

---

concept; for properties, I use italics.

<sup>94</sup> Is the property deflationist at a disadvantage because he must recognize propositions? Well, a desire to avoid propositions certainly seems to motivate some to accept disquotationalism. However, given the

property of truth, but she *deflates* that property: that is, she holds that we should explain truth by reference to simple schemata such as those found in (S1), while holding that when we do so, we are, in effect, saying everything there is to say about the nature of truth.

Of course the schemata employed by the property deflationist cannot be transformed into closed sentences expressing candidate elucidations or analyses of truth, i.e., elucidations or analyses like those that philosophers have provided for such concepts as knowledge or goodness, and this is why property deflationism is a form of deflationism.

## 2.5 A Problem with Idiolect Disquotationalism

Although we have discussed a couple of reasons for favoring property deflationism to disquotationalism, we have not yet considered the most sophisticated and well-developed disquotationalist account, the version known as *idiolect disquotationalism*, which is championed by Vann McGee and (tentatively) by Hartry Field. The purpose of our discussion is to make clear what idiolect deflationism is and to make some problems for it. We do this for two reasons. First, it is of interest to us to compare and contrast the various forms of deflationism that are currently on offer, since we are trying to understand deflationism. Second, it is of interest to us to see which version of deflationism fares the best, so that we can determine which one to employ, when we move on to consider what the deflationist can and should say about meaning. That said, we turn to idiolect disquotationalism.

---

minimalist treatments of propositions that are available, I am inclined to think that the deflationist has little to worry about from them.

The Idiolect Disquotationalist is committed to the following theses:

(ID1) inference-rule deflationism: facts about the nature of truth are exhausted by facts about the meaning of the truth predicate, which is captured by the inference rules

(TI) and (TO)

(TI) From “p” infer “‘p’ is true”,

(TO) From “‘p’ is true” infer “p”;

(ID2) The claim that an utterance  $u$  is true should be read, informally, as  $u$  is true-as-I-understand-it;

(ID3) There are no propositions or inter-subjective meaning relations, though there are standard translations from language to language;

(ID4) For any meaningful utterance,  $u$ , the claim made by uttering “ $u$  is true” is cognitively equivalent (for the person for whom ‘ $u$ ’ is meaningful) to the claim made by an utterance of  $u$  itself; and

(ID5) Cognitive equivalence is to be understood in terms of conceptual or computational role: “for one sentence to be cognitively equivalent to another for a given person is for that person’s inferential rules to license (or, license fairly directly) the inference from either one to the other.” (1994b, p. 405, n.1)<sup>95</sup>

It follows from (ID4) that a speaker can understand an ascription of truth only insofar as she understands the sentence to which truth is being ascribed. It follows from this, together with (ID2) that we cannot ascribe truth to utterances that we do not understand. So, according to the advocate of idiolect disquotationalism, the truth predicate is applicable only to sentences of a person’s idiolect, i.e., only to sentences that she

understands. Hence, the view is a species of *idiolect* disquotationalism. According to Field (1994b, 405), what is registered in (ID2) and (ID4) explicates the claim that truth is at bottom disquotation; hence, the view is a species of idiolect *disquotationism*.

Idiolect disquotationalism has some very interesting properties, not the least of which is that, following Quine (1970), it takes truth to be an *immanent* notion. Furthermore, as just mentioned, per the idiolect disquotationalist, a speaker can apply the truth predicate only to the sentences (or utterances) that she understands, and, when she does, what she says amounts to the claim that the utterance is true-as-she-understands-it. Although this is how Field (1994a,b) puts it, I think it can be stated a bit more perspicuously: the claim—that an utterance of ‘*u* is true’ should be understood as ‘*u* is true-as-I-understand-it’—itself should be recast. An utterance of ‘*u* is true’ should, instead, be understood as ‘*u*-as-I-understand-it is true’. Field’s idea is that truth can be ascribed to utterances that we understand *as we understand them*; in other words, the truth predicate can be applied only to interpreted utterances, and when it is so applied, it is applied to them as they are interpreted.

Putting things in this way makes the view much clearer, I think, for now it is clear why it would be the case that, if I do not understand an utterance, I cannot ascribe truth to it, and it should also be clear why it is that, if I do understand an utterance then, when I ascribe truth to it, what I ascribe truth to is the utterance as that has been interpreted by me, rather than, say, as it was intended by the speaker.

Notice, further, that, given (ID4) and (ID5), it follows that, were I to ascribe truth to my own utterance, the truth ascription would be *cognitively equivalent* (for me) to the

---

<sup>95</sup> A small point: ‘*u* is true’ is cognitively equivalent to ‘*u*’ relative to the utterance of *u*.

utterance itself. In the case of the utterance of another, I first must interpret that person's utterance and then ascribe truth to it as interpreted, in which case, again, the ascription of truth to the utterance, once interpreted, is cognitively equivalent (for me) to the interpreted utterance to which I am ascribing truth.

The idiolect disquotationalist does not take truth to apply to utterances *tout court*, but to interpreted utterance (what Field calls "sentence readings"). We need a way of representing sentence readings and of distinguishing them from ordinary utterances, since this distinction matters to the idiolect disquotationalist. Since sentence readings are just utterances that have undergone interpretation, we can define a function SR which, given an utterance, returns a sentence reading, i.e., an interpretation of the utterance in the person's idiolect. We can then say the following: an utterance is true if what results from an application of SR to it is (disquotationally) true; false if the negation of what results from an application of SR to it is (disquotationally) true; and neither true nor false if SR is undefined for it.<sup>96</sup> An utterance will be *truth evaluable* for a speaker (at a time) only if SR is defined for it.

If, for a person, Stan, and an utterance, u, Stan understands u, then there is a translation of u into Stan's idiolect. There are lots of questions that might arise about Field's use of the notion of an idiolect, the most pressing of which regarding the *size* of a speaker's idiolect—is it infinite? Finite? Countable? What? Does it consist of all of the sentences a speaker *does* understand, or might it consist of the closure under logical implication of that basic set? Is it compositional? These are all questions that might be raised and not all of which have been answered. To make life easy, let's suppose that

idiolects are finite (nothing big turns on this, at least for our purposes), and let's think of a speaker's idiolect (at a specific time) as the set of all sentences that she understands. For further ease, let's suppose that our thinker thinks in a language of thought.

We can understand *understanding* as follows: Stan understands an utterance  $u$  iff there is a sentence,  $p$ , of his language of thought,  $L$ , such that  $SR_L(u) = p$ . We can call the collection of sentences in the speaker's language of thought  $P$ , where, for, Stan, we would say  $P = \{p: \text{Stan understands } p\}$  ( $= \{p | p \text{ is a sentence of Stan's language of thought}\}$ ). Given that, we can then say the following: for an utterance,  $u$ ,

(TRF)  $u$  is true for Stan iff  $(\exists p \in P)(SR_S(u) = p \ \& \ p \text{ is true})$  ;

(FF)  $u$  is false for  $S$  iff  $(\exists p \in P)(SR_S(u) = p \ (\in P) \ \& \ p \text{ is false})$ .

(UF)  $u$  is neither T nor F for  $S$  if  $\sim (\exists p \in P)(SR_S(u) = p)$ .<sup>97</sup>

(DS) " $p$ " is true iff  $p$ .<sup>98</sup>

We should note one thing. If the move to idiolects is to be non-trivial, we should stipulate that for any speaker,  $S$ , there is at least one utterance,  $u_1$ , such that, for  $u_1$ ,  $\sim (\exists p \in P)(SR_S(u_1) = p)$ , i.e., there is at least one utterance that  $S$  does not understand. This stipulation has not seemed to cause any problems for the idiolect disquotationalist, for there seem to be very intuitive ways of handling it. However, as we will see, I think that there is a circumstance under which it might cause trouble. If I am right then there is

---

<sup>96</sup> One result: the idiolect disquotationalist has a *prima facie* problem with truth-value gaps, just as the Horwichian propositional deflationist does.

<sup>97</sup> I leave this as specifying only a sufficient condition because I do not want to get into the issue of truth aptness (To turn it into a necessary and sufficient condition we would have to add only 'and  $u$  is truth apt'). Certain utterances are not truth apt—questions, for example, which, though not assertions, are certainly utterances—though they are certainly meaningful in the sense explained above. Hence, (UF) would be false if we took it to specify a necessary condition on being non-truth evaluable. To make matters easy, and to avoid the question of what truth aptness amounts to, I will leave it as specifying a sufficient condition.

reason to believe that idiolect disquotationalism has an importantly counter-intuitive result.

Return to Stan. Suppose that his idiolect is pretty close to ordinary English but that, as he is mentally deficient in certain ways, it is conceptually limited in certain respects. Suppose, for example, that, though he is thirty years old, his mental age is, and will always be, around 7. So he knows basic arithmetic, but he knows no higher mathematics. Furthermore, as his mental age is 7—as opposed to, for example, 3—he knows that there is quite a bit that he does not, and, indeed, cannot understand. So, for example, he knows that there is much more to math than he understands, but he does not, and cannot, have the foggiest idea as to what it is. Furthermore, though he hasn't the foggiest idea what it is, he knows that lots of other people understand it perfectly well, and that he is special in his inability to grasp it.

To take a simple example, he cannot (even in principle) understand any utterance about calculus—he simply does not have the cognitive ability to grasp the concepts employed in it, and he has not even a partial understanding of most of the concepts that it employs. Suppose that he meets Dr. X, whom he knows to be a university professor of mathematics. Suppose, further, that, as he understands everything that Dr. X says about things that he does understand, as far as he can tell, it appears to be (disquotationally) true, so that, for our speaker, Dr. X appears to be an infallible guide to everything. Suppose, however, that Dr. X is not an infallible guide to everything (who is?), but that, as it happens, her understanding of higher math, and, in particular, calculus, is fatally

---

<sup>98</sup> We could simplify things further, since we are employing something very similar to propositional quantification, which, as is well known, serves much the same logical purpose as truth is expected to serve, but we will not. In fact, we need not elaborate on (TRF) - (UF) an further; they are fine for our purposes.

flawed, so that, as we would judge it, most of what Dr. X believes about calculus is false.

Our speaker conjectures that everything that Dr. X says is true, and utters (35)

(35) Everything Dr. X says is true

In line with the way in which we are taking idiolect disquotationalism, we should understand (35) as (36). The problem with this should be obvious.

(36)  $(\forall u) [(Dr. X \text{ utters } u) \rightarrow ((\exists p \in P_S)(SR_S(u) = p \rightarrow p \text{ is true})]$

*Ex hypothesi*, every sentence reading of Stan's of everything uttered by Dr. X is (disquotationally) true, though, *ex hypothesi*, it is not the case that everything Dr. X utters is true. The point is quite general. If we assume that there are sentences that a person could not understand, perhaps because her conceptual repertoire is, and will remain, too weak, then, per the idiolect disquotationalist, there will be utterances of sentences that are not truth evaluable for that person. According to the idiolect disquotationalist, the function of the truth predicate is to serve as a device of infinite conjunction and disjunction. Hence, if S utters a sentence like (35), what she says is equivalent in some sense to (37).

(35) Everything that Dr. X says is true

(37) If Dr. X says that snow is white then snow is white, and if Dr. X says that snow is green then snow is green, ..., and, ...

As Field says (1994a, especially p. 265), the purpose of an utterance like (35) is to assert an infinite conjunction. Similarly, the purpose of an utterance of (38) is to assert an

(38) What Dr. X just said is true

infinite disjunction, as in (39).

-

(39) Either Dr. X just said that snow is white and snow is white; or Dr. X just said that snow is green and snow is green; or ...

Problems arise when we consider utterances that could not be interpreted by our speaker. We might want to say that, intuitively, (35), is false, since the person to whose utterances truth is being ascribed is a hack; however, if what the speaker asserted was equivalent to the conjunction, (37), which consists only of sentences that he understands, then, since the sentences that he understands are all (disquotationally) true for him, (37), and therefore (35), is true after all. Furthermore, suppose that Stan uttered (38) and that Dr. X uttered some true sentence of calculus, perhaps the only one that she knows. In that case, since, *ex hypothesi*, Stan's idiolect contains no sentence readings of calculus utterances, though we would say that (38) is true, we would say that (39) is false. Hence, if (38) is used to express (39) then, contrary to what we would have thought, it, too, is false. I conclude, therefore, that idiolect disquotationalist makes intuitively wrong predictions about utterances like (35) and (38).

I am trading wholly on our intuitions here, but the point can be made in a different way. Sometimes we utter sentences like (35) and (38) because we believe everything that a person has said, or something that they have just said, only we haven't time enough to run through it all. Sometimes—usually, in the case of something like (35), I would conjecture—we utter sentences like (35) and (38) because we believe certain things about the person who uttered them. If, for example, I form the insane belief that my doorman is god then, since I believe that if there is a god, it is omniscient and not a liar, I will come to believe that everything the doorman believes is true, even though, if the doorman is god, then, as I do believe, some of what he believes may be

wholly incomprehensible to me and, hence, completely uninterpretable by me, even in principle.

If, as it seems, there is nothing wrong with the idea that I can knowingly and successfully indirectly ascribe truth to what a person just said, though, due to some conceptual deficit on my part, what he said will forever resist translation into my idiolect, then there is an important general conclusion we can draw about both truth ascriptions and idiolect disquotationalism. If, in such cases, our linguistic purposes are not thwarted, then it might be that one of the purposes of having blind ascriptions of truth—which, as we know means: one of the reasons for having a truth predicate at all—is that there are circumstances in which we want to ascribe truth to what a person says, though, for various conceptual reasons, we are not in a position to understand what the person has said, and, therefore, again, for conceptual reasons, not in a position to make a direct ascription.<sup>99,100</sup> If this is the case, then it is not clear that the idiolect disquotationalist delivers a notion that plays the logical role that we expect a truth predicate to play.

Field (1994a) considers the alleged problems created for the idiolect disquotationalist by ascriptions of truth to utterances in foreign tongues. He notes three

---

<sup>99</sup> So, for example, we might have occasion to say something like “I cannot understand a word that he says, but I bet that all of it is true,” meaning not that we do not know how what he says translates into our idiolect but that what he says does not translate into our idiolect.

<sup>100</sup> So, to take a simple example, if S, our mentally challenged speaker, overhears two high school students complaining that they do not understand calculus, she might utter (39), intending, thereby, to direct them to Dr. X. Her purposes there do not depend  
 (39) Everything Dr. X says about calculus is true  
 on her understanding any of what Dr. X says about calculus; they depend only on her belief that calculus is the sort of thing that people—herself not included—can understand.

strategies for dealing with them. As we will see, none of them are up to the challenge we have raised.<sup>101</sup>

First, the idiolect disquotationalist might employ a notion of interlinguistic synonymy. On this approach, given a foreign utterance, *u*, we regard “*u* is true” as equivalent to “*u* is synonymous with a sentence of my idiolect that is disquotationally true”. Field is not too attracted to this strategy, but notice that, even if he were partial to it, it could not solve our problem. The reason is that it presupposes that Dr. X’s utterances about calculus are synonymous with the sentences of S’s idiolect, which they are not.

The second strategy is “simply to define what it is for a foreign utterance to be *true relative to a correlation* of it to one of our sentences: a sentence is true relative to a correlation if the sentence of ours correlated with it is true in the purely disquotational sense.” (1994a, p. 273) This option is looser than the first, as it admits of standards of translation without insisting on the *factivity* of translation. The standards of translation that govern the correlations might be highly context-sensitive, interest relative, etc. Does this move help the idiolect disquotationalist with our case? I cannot see that it does. Consider (38). Stan is ascribing truth to an utterance of Dr. X’s, which utterance is described, rather than named directly. Nevertheless, there is no correlation with the sentence so-described to any one of the sentences of his idiolect, and, what’s more, he does not mean to communicate that any of Dr. X’s utterances are, *as he understands*

---

<sup>101</sup> Of course our issue is slightly different from this one: Field is concerned with explaining how the idiolect disquotationalist can handle direct ascriptions of truth to foreign utterances, whereas we are raising a problem posed by the possibility that a sentence or collection of sentences described by a speaker might not be comprehensible to her, and hence, might not be translatable into her idiolect at all.

*them*, true, for he is well aware of the fact that he cannot understand much of what she says.

The final solution that Field considers is that we could incorporate all of those sentences that we do understand into our idiolect, even if, for example they are not sentences of the language that we speak (e.g., English). Hence, if I understand the German sentence “Der Schnee ist weiss”, I can apply truth to it, yielding, “‘Der schnee ist weiss’ is true iff der schnee ist weiss.’ We need go no further with this option, however, for, *ex hypothesi*, our speaker does not understand the sentence(s) to which he is ascribing truth, which is a necessary condition on including a sentence in his idiolect, no matter what language it is couched in. I conclude, therefore, that none of the three strategies can handle the problem that we have raised.

There is one other response to the problem that the idiolect disquotationalist might be inclined to make. First, consider the propositionalist. The propositionalist thinks that only propositions and those things that express propositions are proper candidates for ascriptions of truth and falsity. When we combine this with the claim that only propositions and things that express propositions are, in the appropriate sense, meaningful, it follows that the propositionalist claims that only meaningful entities can be ascribed truth and falsity. If a sequence of sounds is not meaningful, she will say, then there is simply no sense in which we can ascribe truth or falsity to it.

I think that there might be some inclination on the part of the idiolect disquotationalist to argue in the same way: if one genuinely does not understand an utterance, does it really make any sense at all for that person to say that it is true or false? The idiolect disquotationalist takes it that when we ascribe truth or falsity to utterances,

we ascribe it to the utterances as we understand them. Certainly, there must be something right about that. When I ascribe truth to the utterance of another, there is a sense in which I could not but mean that, on my understanding of what the person said, it was true, which explains why something like (40) is complete nonsense.

(40) Your utterance of 'snow is white' is true, but, as I understand it, it is false.

Hence, there seems to be a sense in which, when we ascribe truth to an utterance, we could not but mean that the utterance, as we understand it, is true.

However, on second thought, we should be a bit more careful. It might be the case that, when we ascribe truth to an utterance *directly*, we could not but mean that the utterance, as we understand it, is true. However, this is because when we ascribe truth to an utterance directly—as we do, for example, when we say that it is true that snow is white—we produce a token of the sentence uttered, in which case it is *our* utterance of the sentence to which truth is being ascribed, and our utterance of the sentence, *qua* our utterance, is taken by us as we understand it.

What about when we ascribe truth indirectly, as in the case of (35) or (38)? In that case, we use a description of the sentence (or sentences) that was (were) uttered. One might think that here, as in the case of direct ascriptions, we do not ascribe truth to utterances unless we believe what they say, in which case, again, what matters is the utterance as we understand it, rather than as it is intended, or what have you. However, it is not clear that this is correct. We have pointed to a case in which our speaker believes that some person is god and in which he believes that god is both omniscient and honest. In that case, he believes that whatever god says is true, though he recognizes that, if there is a god, then it might be that much of what god says could not be understood by him.

Hence, when he says that whatever god says is true, he does not mean that on his understanding of what god says, it is true, for he is fully aware of the fact that he has no capacity to understand much of what god says.

Is this case analogous to the propositionalist's case? I cannot see that it is. In that case the thought was that what was said was not meaningful to speaker or hearer, and, hence, not meaningful at all. In our case, the thought is that what god says is fully comprehensible to god—just as what Dr. X says about calculus is fully comprehensible to those with the mental capacity to grasp the concepts; it is just that it is not comprehensible to the person who is ascribing truth to it. We arrive, then, at the nub of the problem: there appear to be circumstances in which we ascribe truth to what is said, based only on facts that we believe about the speaker, and in those cases, the fact that we ourselves do not understand what is said is irrelevant to us. In these sorts of cases, even when we do understand the utterances to which we are ascribing truth, that we understand them plays no role in the explanation for why we ascribed truth to them—i.e., we would have ascribed truth to them even if we did not understand them—in which case it seems appropriate that not understanding them is no bar to ascribing truth to them. The idiolect disquotationalist cannot handle such cases, and, since these cases are type 2 cases, and type 2 cases are the cases that give truth its point, if we are right then the idiolect disquotationalist cannot account for an important function of the truth predicate.

Before we close, I would like to consider a final response to the problem we have posed.<sup>102</sup> Field (1994a), unlike, for example, Horwich, is not interested in providing an account of our ordinary concept of truth; in fact, he doubts that such a thing really exists.

Can't he just bite the bullet and admit that his notion of truth cannot handle the use we are pushing? If he can, he might argue as follows:

You've pointed to a limitation of my view, but this isn't a limitation that worries me much. The claim that we can ascribe truth, directly or indirectly, to any meaningful sentence, whether or not we understand it, presupposes that we can make objective sense of the notion of meaning. But this is just what I deny. So, if I am right, I don't need to adjust my theory to handle your "data"; you need to reset your intuitions, so that they fall in line with my view.

This response is important, for a few reasons. First, it brings to light that disquotationalist presupposition, to the effect that there is no such thing as objective meaning. Of course Field assumes that there is not, but what is the argument? After all, our ordinary practices seem to implicate such meanings.<sup>103</sup> Second, and more important, the 'bite the bullet' strategy is methodologically suspect. As we will discuss, the deflationist leans heavily on our practices—on, among other things, the *purposes* for which we have a truth predicate. To 'bite the bullet' in this case is to adopt a position incompatible with practice. Anyone, especially the deflationist, must give a substantial justification for doing that. As I hope is clear, given the utility of the truth predicate, biting this bullet might be more dangerous than we think.

## 2.6 Field, Substitutional Quantification and Schemata

In the last two sections, we got some reasons for favoring a version of property deflationism over a disquotational view, assuming, of course, that we chose to adopt

---

<sup>102</sup> I would like to thank James Woodbridge for pressing me on this point.

some version of deflationism. I am not certain that the points we raised were conclusive, but, since others have pointed out the problems with disquotationalism (Cf. esp. David, 1994), I will not pursue them further. The issue I would like to turn to next regards the shape of our theory of truth. The options that I will consider are between adopting a theory that consists of an infinite number of axioms, each of which is an instance of (DS), or adopting a single axiom and employing substitutional quantification.<sup>104</sup> As I hope to show, there is a problem with “going substitutional” that makes this a particularly unattractive option and helps to motivate the infinitary view that I think the deflationist should adopt.

At one point (1994, p 259) Hartry Field suggested that a deflationist who wants to capture some of the generalizations that do not follow from the disquotational schema should adopt the principle (TG) as her starting point, employing substitutional (TG)  $(\Pi p)$  (“ $p$ ” is true iff  $p$ ).  
 quantification,<sup>105</sup> where ‘ $(\Pi p)$ ’ binds both occurrences of ‘ $p$ ’. If the deflationist employs the substitutional quantifier,  $(\Pi p)$ , the result will be that (TG) is true iff (DS) is true for

---

<sup>103</sup> In saying this, I am not claiming that our practices support a substantive notion of meaning; they might, but they might also support only a minimalist one.

<sup>104</sup> There is one other option that I will consider, the option of employing schematic letters. We will turn to that option in a bit, for it is the option that I believe that Field should take. I will also consider, on my way to schematic letters, a different version of propositional deflationism, one that I believe that Sosa favors. In the end, I hope to provide some support for the infinitary view.

<sup>105</sup> Field claims that the deflationist needs something like (TG) in order to derive (TFG).

(TFG) For all sentences  $S_1$  and  $S_2$  of our language, ‘ $S_1$  or  $S_2$ ’ is true iff  $S_1$  is true or  $S_2$  is true. Hence, if, as I will show, she really cannot avail herself of (TG), then she cannot justify the claim that ‘or’ obeys the usual truth table. There is another option that she might favor, but that option does not work, either. Although I am discussing the substitutional quantification option in order to show that the deflationist cannot avail herself of it, if both Field and I are right, then the disquotationalist cannot explain everything about truth that needs explaining, which, as we know, would be a problem for the disquotationalist.

all sentences of the language (or, idiolect) for which truth is being defined, i.e., for all substitution instances.

There is a problem with this approach, at least for the deflationist. The approach on which a substitutional quantifier is employed is semantic, not axiomatic, because, in order to explain the substitutional quantifier, we need to invoke a semantics.<sup>106</sup> I think that the disquotationalist will have problems explaining the semantics of the substitutional quantifier, however. Let me explain. In order to explain the semantics governing the substitutional quantifier, what is needed is a notion of truth for a portion of the object language that is free of truth predicates, substitutional quantifiers, or other semantic terms. But this means that what is introduced by the strengthened, substitutional version of (DS)—presumably, a truth predicate for the aforementioned part of the language—is already presupposed in the formulation of the principle used to introduce it.

The deflationist who wants to “go substitutional” has a response to this worry.<sup>107</sup> She can say that the reduction made above—i.e., in the claim that we presuppose exactly what we are trying to introduce—is based on a mistake, the mistake of confusing two different truth predicates. The truth predicates that are confused are the truth predicate for the portion of the object language that is free of semantic terms and substitutional quantifiers—‘true’—and a truth predicate for the metalanguage—TRUE—in which substitutional quantification for the base language is explained. The former applies only to ‘true’-free sentences of the object language, i.e., sentences of the base language. The latter, in terms of which we explain the former, also applies to sentences of the object

language, including sentences containing ‘true.’ I do not think that this response is very helpful, for the deflationist who wants to go substitutional, but it is helpful to us. The reason is that our deflationist is saying that the truth predicate employed in our explanation of substitutional quantification is stronger than the truth predicate that it is being used to explain. Furthermore, and this is the problem, given ‘TRUE’ we can define ‘true’, after all: a sentence, ‘S’ is true iff it is TRUE and it does not contain ‘true’. Hence, in some sense, we are presupposing the notion we aim to explain after all.

Presumably, the deflationist would not want to introduce a substitutional quantifier in the usual semantic way, because she wants to develop a theory of truth and does not want to presuppose a semantics. One thing that she can do is give axioms for propositional quantifiers.<sup>108</sup> Hence, rather than (TG), she might adopt (TP),

(TP)  $(\forall P)(P \text{ is true iff } P)$

Recall that Field claims that we need truth in order to express infinite conjunctions and disjunctions. So, we want to be able to capture generalizations such as (32), which can

(32) Everything Gupta says is true.

be represented formally as (33). Armed with (TP), we can express what is captured in

(33)  $(\forall x)(\text{Gupta says } (x) \rightarrow \text{True } (x))$

(33) using (41), which does not invoke the truth predicate at all.

(41)  $(\forall P)(\text{Gupta says } (P) \rightarrow P)$

Of course Field wants to avoid quantifying over propositions, so he would not go for something like (TP). However, even if he would not, someone else might, so we have

---

<sup>106</sup> I would like to thank Arnie Koslow for discussions on this point.

<sup>107</sup> I would like to thank Arnie Koslow for bringing this response to my attention.

<sup>108</sup> A suggestion similar to this one is floated in Sosa, 1993.

an interest in seeing if it will work. As it happens, it will not. In effect, if we can avail ourselves of (41), then we do not need a truth predicate after all. Hence, if we introduce truth using something like propositional quantification, our approach will be, in a sense, circular: we will have introduced truth by means of a device that can serve exactly the same purpose.

At any rate, as previously mentioned, Field would not go for (TP). What Field would prefer, since full-blown substitutional quantification seems out of the question, is to “go schematic”, which means that, rather than introducing a substitutional quantifier, he would introduce schematic letters and makes use of truth schemata, such as (DS).

There seem to be two problems with this proposal, however. First, consider how, armed with the schematic letters, we might frame (33). Presumably, we would advert to something like (42).

(42) Gupta says  $P \rightarrow P$  is true,

where ‘P’ is a schematic letter. Notice, however, that in (42), ‘true’ is superfluous. We could do as well with (43). Let’s leave that worry to the side, though. Notice that if we

(43) Gupta says  $P \rightarrow P$ .

go with the schematic version, forsaking the substitutional one, then we are back where we started. If Field was right that we could not capture all of the generalizations that need capturing by means of (DS), then the schematic approach is insufficient. What we really need is not something like (DS) or (42), but, rather, a universally quantified correlate to those; however, if our arguments regarding substitutional quantifier are correct, there is good reason for thinking that this is not something that we can get.

I am sure that Field can make some response to these arguments, but, until a good one is presented, and given all of the other worries we have raised about disquotationalism, it seems that if one is going to opt for deflationism, one should opt for some form of property deflationism, such as some version of the infinitary, axiomatic version of propositional deflationism that we have discussed.

## 2.7 More on Propositional Deflationism

In what follows, I will work with a version of propositional deflationism that is modeled on Horwich's minimalist notion of truth. I should note, however, that the arguments that I provide hold as well for other version of deflationism, including Field's idiolect disquotationalism.<sup>109</sup> The version of deflationism that I will now sketch—a form of generalized, or thorough-going deflationism—is the most liberal view compatible with the claim that all of the facts about truth that need explaining can be explained by reference to the schemata listed in (S1), together with uncontroversial facts and theories. Following Horwich (1990, 1998), we will make use of so-called *propositional forms*. As Horwich lays it out, correlated with each schema in (S1), there is a propositional form, a function from entities to propositions. Hence, we have (S2):

- (ESP) <“p” is true iff p> ---- “p” is true iff p  
 (EPP) <<p> is true iff p> ---- The proposition that p (= <p>) is true iff p  
 (EBP) <A belief that p is true iff p> ---- A belief that P is true iff p  
 (EUP) <An utterance of “p” is true iff p>--- An utterance of “p” is true iff p

To be a proposition of a certain form is to be the value of the propositional form for some argument. So, to take a simple example, <p v q> is a function, DIS, which yields the

proposition, DIS (p,q), given arguments p and q;  $\langle \sim p \rangle$  is also a function, NEG: given, as argument, the proposition, p, NEG returns a proposition NEG(p), the value of NEG for the proposition, p. Certain propositions *decompose* into others. To take an example, consider an instance of EPP, e.g., (43).

(43)  $\langle \langle \text{snow is white} \rangle \text{ is true iff snow is white} \rangle$

EPP is decomposable into the functions IFF and T, so that we can read (43) as (44)

(44) IFF[T( $\langle \text{snow is white} \rangle$ ), snow is white],

where IFF is construed on analogy with NEG and DIS, and where T is the function that takes a potential truth bearer and returns the proposition that it is true. So,  $\text{EPP}(\langle \text{snow is white} \rangle) = (44) = (43)$ .

Furthermore, a proposition is of the form  $\langle \langle p \rangle \text{ is true} \rangle$  just in case it is the value of T for some proposition, p. If we introduce a deflationary notion of what it is for a sentence or belief to express a proposition (N.B., as Horwich (1998b) does), we can even explain sentence truth in terms of propositional truth, thereby mimicking more substantive accounts of truth. In order to do this, we introduce the expression function, (EXP). (EXP) is a function from a sentence to the proposition it expresses,<sup>110</sup> i.e.,  $\text{EXP}(\text{"p"}) = \langle p \rangle$ .

Armed with our deflationary expression (or, meaning) function—deflationary because what it is to express a proposition is being explained by reference only to a deflationary principle, (EXP)—we can explain the instance of (ES) for “snow is white”; i.e., (5),

(5) “snow is white” is true iff snow is white

---

<sup>109</sup> See previous section for a reason to favor propositional deflationism over idiolect disquotationalism.

<sup>110</sup> It has been suggested to me that (EXP) must be a function from sentences and languages to propositions. Although this point is technically important, we can ignore it for present purposes.

as follows, showing, thereby, that (ES) can be reduced to, or decomposed in terms of,

- (a)  $EXP(\text{"snow is white"}) = \langle \text{snow is white} \rangle$ ;
- (b)  $EP(\langle \text{snow is white} \rangle) = \langle \langle \text{snow is white} \rangle \text{ is true iff snow is white} \rangle$ ; hence,
- (c)  $ES = EP(EXP(\text{"snow is white"}))$ , .

(EP) and (EXP). Similar, though slightly more complicated, accounts can be given for (EB) and (EU). We can then follow Horwich and identify our deflationist theory of truth by reference to the aforementioned schematic forms: our theory will consist of an infinite collection of propositions of those types. Of course we can also define truth for non-propositional truth bearers in terms of truth for the propositions they express, in which case our theory of truth would consist of all and only acceptable instance of (EPP), together with the relevant instances of (EXP), etc.<sup>111</sup>

## 2.8 What is it to Explain Truth? (One view)

Why move to the propositional forms? The move to propositional forms is important, for one thing, because it allows our deflationist to explain truth.<sup>112</sup> Without something like the propositional forms, our deflationist has only the schemata listed in (S1), and the schemata do not provide the deflationist with what she needs if she is to

---

<sup>111</sup> There are other options available to our deflationist, of course. She might prefer a finite theory, in which case she would take her theory to consist of propositions of the following sort:

For all sentences,  $s$ ,  $EXP(s)$  is materially equivalent to  $TRUE(s)$

For all propositions,  $p$ ,  $p$  is materially equivalent to  $TRUE(p)$

For all beliefs,  $b$ , if  $b$  is a belief that  $p$  then  $p$  is materially equivalent to  $TRUE(b)$ .

For all utterances  $u$ , if  $u$  is an utterance of  $s$ , then  $EXP(s)$  is materially equivalent to  $TRUE(u)$ , where  $TRUE$  and  $EXP$  are functions from potential truth bearers to propositions.  $TRUE$  is a function from a potential truth bearer to the proposition that it is true, and  $EXP$  is a function from a potential non-propositional truth bearer to the proposition it expresses.

What he would also need in the finite theory is a principle of material implication (or something equivalent) that he could use to derive instances of the various equivalence schemata, as well as certain other unproblematic premises, such as the proposition that water is  $H_2O$  is the value of  $EXP$  for the sentence 'water is  $H_2O$ ,' and the proposition that 'water is  $H_2O$ ' is true is the value of  $TRUE$  for the sentence 'water is  $H_2O$ '. Cf. Sosa, 1993.

<sup>112</sup> In saying this, I do not mean to imply that this is the only method that the deflationist can use to generate his theory of truth, only that it is one method.

explain all of the fact about truth that need explaining. In order to explain all of the facts about truth that need explaining, we need a *theory* of truth, a collection of propositions, sentences, claims, or what have you, that purports to explain truth (A good theory does more than purport, of course), and the various schemata, though they might provide a recipe for a theory, do not amount to a *theory* of truth.<sup>113</sup> The propositional forms deliver a theory of truth, understood, as previously mentioned, as an infinite collection of instances of the various schemata. Whether or not it is a good theory is precisely what will be at issue later on.

So, we move to propositional forms—or, at any rate, Horwich (1990, 1998) seems to have—because the schemata are insufficient to the explanatory needs of the deflationist, but what exactly are the explanatory needs of the deflationist? There are two questions that have to be separated out:

(Q1) What is it to ascribe to truth an explanatory role?

which I take to mean, at least in one application, What (theoretical or practical) role is played by truth ascriptions? That is, what are we doing when we ascribe truth? We will attack this issue in the next chapter. The other question is:

(Q2) What is it to explain truth?

I take (Q2) to mean something like this: given that we know that (45) is true, how can (45) <snow is white>

we explain that fact, and what is it to explain that fact? More to the point, we might wonder, In what sense do propositions of the form (ESP), (EPP), (EBP) and (EUP) explain truth? These are the questions with which we will be concerned in what follows.

---

<sup>113</sup> I would like to thank Chris Gauker for very useful conversations on the problems the deflationist has

Our answer to these questions will be important, for, as we will see, there is good reason to believe that there are facts about truth that cannot be explained by our deflationist.

What we want to explain are propositions of the form  $\langle\langle p \rangle \text{ is true}\rangle$ ,  $\langle\text{"p"} \text{ is true}\rangle$ , etc. Following Horwich, our deflationist claims that what explains such propositions are propositions of the form (EPP), (ESP), etc. That is, a proposition of the form (EPP) explains a proposition of the form (TP) in virtue of the fact that it cites as

(EPP)  $\langle\langle p \rangle \text{ is true iff } p\rangle$

(TP)  $\langle\langle p \rangle \text{ is true}\rangle$

explanans of (TP) what are, in fact, its explanans—what explains the fact that  $\langle\text{snow is white}\rangle$  is true is the fact that snow is white; that is, it is *because*  $\langle\text{snow is white}\rangle$  is the explanans of  $\langle\langle\text{snow is white}\rangle \text{ is true}\rangle$  that we can say that an instance of (EPP) explains

an instance of (TP). Let's take another example. Suppose that we want to explain the proposition, (46); that is, as we will see in a few moments, we want to say what would

(46)  $\langle\langle\text{fire is wet}\rangle \text{ is true}\rangle$

explain this proposition, irrespective of its actual truth value. Given that proposition as argument, the deflationist's theory yields an appropriate instance of (EPP), (47).

(47)  $\langle\langle\text{Fire is wet}\rangle \text{ is true iff fire is wet}\rangle$ ,

which, the deflationist claims, is an explanation of the proposition (46). She claims that (47) is an explanation of our target proposition, (46), because (47) cites as explanans of that proposition what is its explanans, namely, (48).

(48)  $\langle\text{fire is wet}\rangle$ .

---

with moving from schemata to theories.

I have intentionally chosen as our model case a false proposition, in order to bring out an aspect of the notion of explanation that our deflationist is employing that seems to be puzzling. The notion of explanation that she is employing seems to be the notion of explanation that figures in philosophical analyses—the relation between explicandum and explicans is analogous to the relation thought to hold between analysandum and analysans; i.e., it is the notion of explanation employed by the epistemologist, when she tells us that knowledge can be explained in terms of true justified belief, or, better, that a person knows that *p* iff she has a true justified belief that *p*; and it is the notion of explanation that is employed by the Utilitarian ethicist, when she tells us that goodness can be explained in terms of the production of the greatest happiness for the greatest number of people, or, better, that an act is good iff it produces the GHGN.

In short, the explanation relation that deflationists like Horwich employ is an a priori relation holding among propositions, and, as regards truth, it can be captured either by schematic forms like  $\langle T(\text{"p"}) \text{ iff } \langle p \rangle \rangle$ , or, if we like, by schemata of the form  $\langle \langle p \rangle \text{ explains } T(\text{"p"}) \rangle$ .<sup>114</sup> This notion of explanation is non-factive, which is why we can explain  $\langle \langle \text{fire is wet} \rangle \text{ is true} \rangle$  by citing  $\langle \text{fire is wet} \rangle$ , though neither proposition is true. It does sound a bit odd to explain a false proposition; however, this oddness is a consequence of the fact that this notion of explanation should probably be paraphrased as ‘would explain’, rather than as ‘explains—  $\langle \text{fire is wet} \rangle$  would explain  $\langle \langle \text{fire is wet} \rangle \text{ is true} \rangle$ , just as Allison’s having a justified true belief about *p* would explain her knowing that *p*, even though, as it happens, everything she believes is false.’<sup>115</sup>

<sup>114</sup> This can be decomposed as follows: [EXPL ( $\langle \text{fire is wet} \rangle$ ,  $T(\langle \text{fire is wet} \rangle)$ )].

<sup>115</sup> I would like to thank Jonathan Adler for very helpful discussion and Matt McGrath for very helpful email correspondence on the points in this section.

Finally, the philosophical explanation relation implies the material equivalence of explanans and explicandum, as is the case with analyses. So, if F-ness is explained by (or, analyzed in terms of) G-ness, then  $x$  is F iff  $x$  is G. This means that if the biconditionals that make up, or are entailed by, the deflationist's theory of truth are understood to capture explanation relations, then the theory implies the material equivalence of  $S$  and  $S$  is true, for all potential truth bearers,  $S$ , thereby satisfying Tarski's famous condition of adequacy, according to which a theory of truth should imply all non-paradoxical equivalences of the form  $S$  is true iff  $S$ .

## 2.9 Semanticalism

In the last section, I discussed briefly how our deflationist might understand what it is to explain truth. As we will see in chapter 4, there will be certain facts about truth that do not seem explicable on the deflationist's theory. In the next chapter, we will turn to (Q1), and we will investigate charges made against deflationism, to the effect that the theory cannot account for the explanatory role of truth (and truth conditions), which role requires the provision of a substantive theory of truth. I would like to close this chapter with a return to the topic of semanticalism, for this issue, though it is of central importance to the deflationist, has not received much attention in the literature.

Recall that, as we mentioned previously, Field (1972) criticized Tarski's account of truth for being incomplete: as it stood, Tarski seemed committed to semanticalism, which would render his program unacceptable to the physicalists. Furthermore, since one of Tarski's goals was the provision of a physicalistically and methodologically acceptable account of truth, it would, thereby, undermine his philosophical and methodological

(though not his technical) goals. Leeds (1978) responded to Field by pointing out that if truth (or primitive denotation) is understood *disquotationally* then it cannot function as a centrally important, causally-explanatory theoretical notion, and if it does not so function, then it does not have to live up to the canons of good explanation for such notions, in which case it can be left unreduced.

I am a bit unclear on Leeds' claim, and for the following reason. In his 1978 article (see especially pg. 122), Leeds suggests that the deflationist can avail herself of Tarski's methods and convert the implicit definition of truth that she has into an explicit definition of truth. Although I do not want to press too hard on this point, I do not see this as plausible. Tarski's methods required stronger machinery than the deflationist allows herself (recall that Tarski's method for defining truth for a first-order language involved the use of second-order logic). Moreover, as Tarski himself showed, we cannot, in L, explicitly define truth for our language, L, which means that if we want an explicit definition of truth, we have to give it in a metalanguage, using metatheoretical resources.

Now, as I understand it, it is a virtue of deflationism that the deflationist does not retreat to a metalanguage, and avail herself of a substantial metatheory. What makes deflationism attractive (one thing, at least) is that all that the deflationist avails herself of is what is absolutely necessary for understanding truth, which means, in effect, the truth predicate and the principles that govern it.<sup>116</sup> But if this is the case, then the deflationist cannot transform her implicit definition into an explicit definition, in which case, it does not seem that Leeds is correct.

---

<sup>116</sup> There is a question about how much machinery the deflationist can/should avail herself of. Field (in conversation) suggested that, if necessary, she should avail herself of all of the metatheoretical tools that

I do not know how important this last point is, but, assuming it is correct, this means that, as noted previously, the deflationist is committed to the claim that if there is a property of truth, it is both primitive and undefined. We noted previously that this made it look as though the deflationist was committed to semanticalism, since semanticalism is the view that truth is a primitive, undefined notion. We noted that there was reason to believe that the deflationist was not committed to semanticalism, because there was reason to believe that semanticalism was the belief that truth is a primitive, undefined, *explanatory* property, and, as we know, the deflationist would sooner give up her view than accept that.

Although we concluded that the deflationist was not committed to (a vicious strain of) semanticalism, now that we have distinguished between the two most prominent deflationist views, we might go back and consider how each of them avoids semanticalism. As we will see, doing so will give us even more insight into the difference between an inflationist and a deflationist view of truth. In what follows, then, I will consider the question of how an idiolect disquotationalist might avoid semanticalism and I will consider the question of how a property deflationist might avoid semanticalism.

### **2.9.1 Idiolect Disquotationalism and Semanticalism Revisited**

In order to see how the idiolect disquotationalist avoids semanticalism, we must say a bit more about her view. For this, we turn to Hartry Field. According to Field (1994a,b), as previously mentioned, the disquotationalist holds that instances of the

---

Tarski employed. Halbach (1999) is more conservative, holding that the deflationist may not avail herself of any metatheoretical tools.

disquotational schema hold in virtue of the meaning of the word 'true': instances of (DS) (for the sentences I understand) are logically true (for me), in virtue of the fact that a sentence, *S*, and the sentence to the effect that it is true, are cognitively equivalent, which is to say that my inferential rules license (fairly directly) the inference from one to the other and back again. Hence, Field claims that instances of (DS) are logically true in virtue of their meaning.

Recall that for Field the relevant notion of meaning takes the meaning of a sentence (for a person) to consist in, at least, the role that that sentence (or its internal analogue) has in that person's computational psychology. Hence, Field's claim that instances of (DS) are true in virtue of the meaning of 'true' amounts to the claim that our inferential rules sanction the move from the left-hand side to the right-hand side and back again. It is not clear (to me) in what sense this makes instances of (DS) "logically true",<sup>117</sup> for it is not clear in what sense, for Field, they are true at all. Of course Field holds that the fact that our inferential rules license the inference from "*p*" to "*p*" is true and back again is *explanatorily fundamental*, meaning that it cannot be explained by appeal to any further facts about us or our procedures. Hence, it looks as though we can say that for Field and the idiolect disquotationalists, instances of (DS) (for the sentences we understand) are explanatorily fundamental, as well, which means that, like electromagnetic phenomena, the instances of (DS) cannot be explained by reference to further facts or theories. So, although the idiolect disquotationalist denies that the fact

---

<sup>117</sup> When Field says that they are logically true (i.e., in his 1994b, p. 405, n.1), he, too, keeps 'logically true' in scare quotes. Does this mean that he, too, wonders about the sense in which they are *true*? Or does he worry about the sense in which they are *logically* true?

that the instances of (DS) hold is explanatory of anything, it looks as though she must accept that they are primitive, basic facts.

Although this issue has not been discussed (to my knowledge), I think that the claim that Field and the idiolect disquotationalists are committed to even this seemingly innocuous version of semanticalism is based on a mistaken view of how the idiolect disquotationalist views, or should view, instances of (DS). I believe that the idiolect disquotationalist should deny that we need to explain why instances of (DS) hold because she should deny that there is any interesting sense in which they do hold. Her explanation for why we have a notion of truth, and, indeed, for the reason that the instances of (DS) are important, does not require her to claim that there is any real sense to the claim that, as Field might describe it, an utterance of (49) holds, or is true.

(49) It is a fact that *snow is white* is true *iff* snow is white.

If the idiolect disquotationalist denies that sentences like (49) are true, and, given Godel's second incompleteness theorem, we have reason to believe that she will, then her need to explain why instances of (DS) hold is obviated, for if she denies that they are true, there is no interesting sense—indeed, there is no sense—in which they do hold. Of course we can explain why we accept them—by reference to the conceptual role of 'true'—but that does not entail that they are factual claims. Hence, I think that the idiolect disquotationalist can avoid semanticalism, but I do not see that she does so in the way that, or for the reason that, Leeds might have suggested.

We can approach this in a slightly different way. According to Quine, "By calling the sentence ["snow is white"] true, we call snow white." (1970, p. 12) For this reason, Quine tells us that, if we want to know what makes the sentence "snow is white" is true,

we ought to learn more about snow. But this raises an interesting question. If, in saying that the sentence, ‘snow is white’, is true, we call snow white, and if our interest in knowing what makes ‘snow is white’ true should lead us to study the properties of snow responsible for its reflecting and refracting light as it does, what should we say about (50)?

(50) “snow is white” is true iff snow is white

What are we saying when we utter that sentence? If it is true, what do we have to learn more about to know what makes it true? Truth? But there is nothing more to learn about truth? The sentence “snow is white”? I think that the answer that the disquotationalist will give is clear: though you cannot learn more about (50), by learning more about truth, there being nothing more to learn about truth than is captured in (50) and in sentences like it, you can learn more about why we accept (50); about what role sentences like (50) play in allowing us to formulate certain generalizations that we could not otherwise formulate; about what role sentences like (50) play with respect to our grasp of the concept of truth; etc. I believe that this provides us with a reason for believing that, for the disquotationalists, talk about truth is *non-factual*—if you want to learn more about (50), the disquotationalist will not steer you in the direction of any facts, except of course facts about how we (she?) use the truth predicate.

I believe that the deflationist—i.e., the property deflationist or the idiolect disquotationalist—would respond to my question about the truth of (50) by claiming that it is not an interesting question: we have no reason to concern ourselves with whether or not the instances of the disquotational schema are true; furthermore, we have good reason for thinking that your question about the truth of (50) is not an interesting one. It

is not an interesting question because nothing about the function of the truth predicate suggests that it need be answered—it is not an interesting question because it does not matter one way or the other to us whether or not the instances of (DS) are true. What matters is that we accept them, and that it is because we accept them that the truth predicate is able to serve us in the way that it is.

This is fair enough, but I think that it might be a bit weaker than the truth (so to speak). The truth is that if the instances of (DS) are axioms of the deflationist's theory of truth (or, just: if they are explanatorily basic), then the deflationist simply cannot entertain the question of whether or not they are true. In order to establish that the axioms of the theory are true, she would need to appeal to some other theory, and it is central to the deflationist's position that there is no other theory that we can appeal to. Rather than claiming that the axioms are true, but that this cannot be proved, which I do not think that the deflationist would want to do anyway, given what we have just said, it is clear that the deflationist has good reason for denying that the instances of (DS) are true, and, of course, if she denies that they are true, if, for example, she interprets the semantic portion of our language instrumentally, then she is not threatened by the threat of semanticalism (though she will have to explain how the factual/nonfactual distinction is to be drawn).

### **2.9.2 Property Deflationism and Semanticalism**

Horwich agrees that instances of (DS) are conceptually basic; in fact, as previously mentioned, he (1997, pp. 95-967) takes the central idea behind deflationism to be that we accept instances of (DS) in the absence of either inference or argument. What

relevance does this have to the property deflationist's program? After all, given the distinction between conceptual basicness and explanatory fundamentality, it is obvious that something can be conceptually basic though it is not explanatorily fundamental.<sup>118</sup> I believe that, for the propositional deflationist, the claim that instances of (DS) are conceptually basic is absolutely crucial, but for a strategic, rather than a philosophical (or technical), reason.<sup>119</sup> Horwich purports to be capturing our ordinary notion of truth. If, on our ordinary conception of truth, instances of (DS) are accepted without inference or argument, then the substantivist needs a reason for claiming that we must explain why they hold. Hence, Horwich's claim that the instances of (DS) are conceptually basic ensures that the burden of proof shifts to the substantivists.

The obvious way to establish that instances of (DS), though conceptually basic, are not explanatorily basic, is to provide some reason for thinking that they need to be explained, and the only plausible reason for thinking that they need to be explained is that they play some explanatory role that requires that they be explained (in a more-than-deflationary way). Since the explanatory role that the deflationists accept—the logical role noted by Quine (1953, 1970)—does not require a more-than-deflationary account of truth, Horwich has a *prima facie* case against the substantivist.

Our immediate question regards semanticalism, however: is the property deflationist committed to semanticalism? The short answer is that property deflationism is committed to semanticalism. The property deflationist accepts that there is a property of truth (admittedly, it is not an explanatory property), and she accepts that it is both

---

<sup>118</sup> Indeed, as we will see, instances of the schema, POSS  $\sim$ (“S” is true iff S), are conceptually basic, though, per Field, they are not even true.

primitive and indefinable. Hence, her view is not unlike the highly inflationist primitivist view once championed by Russell and Moore.<sup>120</sup> On that view, ‘true’, like ‘good’, expresses a primitive, indefinable property.

In an unpublished paper, Horwich acknowledges the similarity between his view and the old primitivist line,<sup>121</sup> but he attempts to differentiate the positions by claiming that truth, on Moore and Russell’s view, is both primitive *and* explanatory. I have gone back and read everything that I could find that Moore and Russell have written on the primitivist view of truth—there was not much<sup>122</sup>—and nowhere do they claim that truth is an explanatory property. I suspect that what Horwich meant was that, on their view, truth is an explanatory property in the sense that if it could be shown that truth *did* have an explanatory role to play, that would not lead them to give up the primitivist line, whereas, if it could be shown to Horwich (or some other property deflationist) that truth played a substantial explanatory role, that would provide a reason for giving up the primitivist view. If this is what would distinguish Horwich’s view from theirs, then there is an important sense in which the views are different, even if, as things stand, they agree that truth is a primitive, indefinable property.

As I hope is clear, then, the property deflationist is committed to some form of semanticalism. However, it is not clear that the form of semanticalism to which the property deflationist appears committed is in any way threatening. I will let the reader decide for himself whether or not this is the case.

---

<sup>119</sup> What follows is closely related to a discussion from earlier in this chapter.

<sup>120</sup> See Richard Cartwright (1987), for a fascinating discussion of this view. Sosa (1993) attempts to reanimate it, but not to much success, I believe.

<sup>121</sup> This is a draft of what would later become part of his book, *Meaning*.

<sup>122</sup> Moore: 1899, 1901, 1911; Russell: 1903, 1904, 1910.

Before I close this chapter, I would like to note that, whether the deflationist adopts disquotationalism or property deflationism, there is an important sense in which she is a non-factualist about truth. There are numerous ways in which the factualist/non-factualist distinction can be drawn. The two most famous ways of characterizing the distinction are in terms of the possession of truth conditions (“the semantic distinction”) and in terms of the expression of beliefs (“the psychological distinction”). On the first way of drawing the distinction, i.e., the way favored by philosophers like Paul Boghossian (Cf. Boghossian, 1990,1993) and Frank Jackson, Graham Oppy, and Michael Smith (Cf., Graham, Oppy and Smith, 1994), the question of whether or not ascriptions of truth are to be understood as factual or non-factual turns on whether or not they themselves have truth conditions, i.e., on whether or not they are *truth apt*.<sup>123</sup> On the second way of drawing the distinction, the question of whether or not an ascription of truth is factual or non-factual turns on whether or not truth-ascribing utterances count as expressions of belief.<sup>124</sup>

Although I think that there is something intuitive in each of these characterizations, and although it is not my interest to join the debate over the proper way to understand the distinction, I would like to add a more general way of drawing it.<sup>125</sup> As I will show, this way of drawing the distinction makes both the property deflationist and the idiolect disquotationalist non-factualists about truth without, at the same time, undermining the property deflationist’s claim that there is a property of truth, though not

---

<sup>123</sup> I was introduced to this term by the paper by Graham, Oppy, and Smith. I have no idea where it originates.

<sup>124</sup> Presumably, this is because beliefs, as standardly understood, have truth conditions. Are the two ways of drawing the distinction interestingly different?

a substantive, naturalistic, complex, or causal-explanatory one. I believe that a more general way to draw the distinction between factual and non-factual concepts<sup>126</sup> is in terms of their *function*. The function of a factual concept (or predicate), and of the utterances in a factualist domain of discourse, is to describe and characterize, whereas the function of non-factualist concepts is not to describe and categorize.

Although I recognize that this characterization is vague, and although there is much more that can be said,<sup>127</sup> it is clear that, if a concept is factualist iff it has the function of describing and characterizing,<sup>128</sup> then truth is not a factualist concept and truth talk is non-factualist, at least according to both the property deflationist and the idiolect disquotationalist. Notice, further, that this provides us with another way of understanding how the deflationist turns the tables on her opponents: if we add that we should only expect a reductive analysis of factualist notions, we can say that, unless the substantivist furnishes an explanatory role for truth to play then, assuming she accepts what we have said about both the factualist/non-factualist distinction and about the conditions under which we should (or, shouldn't) expect a reductive analysis, she should not expect a substantive account of the notion of truth.

### **Chapter 3: Explanatory Liaisons**

Now that we know, in at least rough outline, the shape of some of the most important contemporary deflationist theories of truth, and, in particular, now that we

---

<sup>125</sup> This way of drawing the distinction owes a lot to the work of Huw Price, though I do not know if he would endorse it.

<sup>126</sup> I intend this characterization to carry over to the domain of discourse built around these concepts.

<sup>127</sup> And that I hope to say, though not in this dissertation.

<sup>128</sup> I assume that what it is to “describe and characterize” is reasonably clear. If it is not, then an alternative account of the function may have to be given. This point is orthogonal to our purposes, however.

know what sort of explanation of truth deflationists like Horwich are willing to give, we can ask whether or not a deflationist position—property deflationist or idiolect disquotationalist—is adequate to our purposes. For criteria of adequacy, we turn to Horwich, who writes “The primary test of minimalism (and any other) theory is its capacity to accommodate the phenomena in its domain. That is to say, if our theory is a good one, it will be able to account for all of the facts about truth.” (1990, p. 22) So, per Horwich, to be adequate, a deflationary theory of truth must provide us with the materials necessary to explain all of the facts about truth that need explaining, which means that what we must do is consider the various facts about truth that need explaining, and see if our deflationist can accommodate them.<sup>129</sup>

### 3.1 Explaining Truth and the Explanatory Role of Truth

Hartry Field (1986) argues that if truth conditions have a certain sort of explanatory role to play, then we would need to explain how it is that sentences or utterances get their truth conditions, and he thinks that if we need to explain how it is that sentences or utterances get their truth conditions—alternatively: what it is for a sentence or utterance to have the truth condition that it has—then we will need a substantive, non-deflationary account of truth. The reason is that, for Field, the call for an explanation of truth conditions just is the call for a non-deflationary theory of truth, for, as he sees it, a theory of truth might properly be called a theory of truth conditions. It should be obvious why this is the case: suppose that we needed to explain what it is for a sentence, *S*, to

---

<sup>129</sup> As I prefer propositional deflationism, i.e., for reasons already given, because it is more liberal, and for other reasons that are irrelevant to our present concerns, I will employ it in what follows, though nothing in the arguments turns on doing this. Of course I will consider disquotationalism where it is appropriate to do so, as well.

have the truth condition that it has, i.e., *p*. In that case, we would be looking for some feature of *S*—call it feature *F*—such that *S* has the truth condition that *p* if and only if *S* is *F*. If we found such an *F* then, like it or not, we would be on our way to a non-deflationary theory of truth: we would know that one way for a sentence to be true is for it to have *F* and for it to be the case that *p*.

If Field is right, and I suspect that he is, then, when this claim made by Field is combined with Horwich's anti-substantivist strategy (discussed in the previous section), it should be clear that an important part of the debate between deflationists and their opponents turns on the explanatory role of truth. If (sentence or utterance) truth—alternatively: the fact that a sentence or utterance has the truth condition that it has—is explanatory of anything, then it, too, will need to be explained, and the claim that it needs explaining is the linchpin of non-deflationary accounts of truth. Furthermore, the claim that there is nothing that truth explains has played an important role in the deflationist's argument, for if truth is not an explanatory property, then it is likely that it is not the sort of property that admits of a reductive analysis, which means that it is most likely not a property that needs the sort of explanations that substantivists have attempted to give of it. That is, on at least one view, unless truth is explanatory then it is not in need of explanation, and if it is not in need of explanation then we have no reason to expect there to be an explanation of it.

The question of whether or not truth is explanatory has created lots of heat, but not much light. In general, deflationists have been unimpressed with the arguments floated by the substantivists and the substantivists have been unmoved by the deflationist's contention that truth is explanatorily otiose. I believe that various issues

about truth and explanation have been thrown together and I would like to separate them out. After they are separated, I think that we will see that there are really two important issues that must be addressed. As I will argue, even if the deflationist is correct about the explanatory role of truth, it does not follow that she is clear of explanation-relation problems.

### 3.2 Explanatory Role?

There are three issues that arise with respect to the explanatory role of truth. First, why care about it? Second, what is it? And third, given the actual explanatory role of truth, what view does it support? In what follows, I will attempt to answer all of those questions. However, before we turn to them, we should consider the question What does it mean to say that truth has an explanatory role?

To talk about the explanatory role of a notion is, obviously, to talk about the role played by that notion in one or another of our explanatory practices. The claim that truth is an explanatory property, or that truth conditions have an explanatory role, then, is the claim that the property of truth (or the possession by an utterance or belief state of truth conditions) plays a role in one or another of our explanatory practices. One of the most common place to look for an explanatory role for truth is in psychological explanation, e.g., in so-called *success* explanations—our explanation of successful behavior.

Put very briefly, and restricting ourselves to the success of individual action,<sup>130</sup> numerous substantivists (e.g., Putnam, 1978, Field, 1986, Devitt, 1994) have argued that

---

<sup>130</sup> The reason for this is that it combines two important areas in which truth has been thought to play an important explanatory role: success explanations and psychological explanations. The explanations of successful action—in particular, the explanation of successful *intelligent* action—combines considerations from these two domains.

we need to invoke the property of truth in order to explain *patterns* of success and failure. In particular, they note that some beliefs are conducive to success (understood, for example, as the satisfaction of desires), while others are not, and they ask after the difference between the two sorts of beliefs: what distinguishes those beliefs conducive to desire fulfillment from those beliefs the possession of which by an agent tends not to lead to desire fulfillment, i.e., those that tend to lead to the frustration of that agent's desires? Truth comes in at this point, for it is argued that a plausible role for truth in explaining an agent's behavior is as follows: "An organism's belief that is true is conducive to the fulfillment of its desires, and one that is false is conducive to their frustrations." (Devitt, 1994, p. 99)

Leaving to the side the plausibility of the claim, it is clear enough what explanatory role is being imputed to truth: the substantivist wants to say that certain beliefs are conducive to success *because* they are true, while other beliefs lead to the frustration of our desires *because* they are false. There are numerous ways to make this claim more precise. Although I will be brief, the way that I favor takes our talk of explanation to be cashed out in terms of explanation relations holding among propositions. On this view, the claim that Allison succeeds because her beliefs are true—the claim that the fact that her beliefs are true explains her success—can be understood as the claim that the proposition that Allison's beliefs are true bears the explanation relation to the proposition that she succeeds.<sup>131</sup> If the substantivist provides a sufficient reason for believing that there is an explanatory relation between (e.g.) the proposition

---

<sup>131</sup> I have not, and will not, describe the characteristics of the explanation relation, as there are various proposals that have been put forward and there would be no point to going with one over the others.

that Allison's belief's are true<sup>132</sup> and something else—the proposition that she succeeds at some task, for example—then she provides us with a compelling reason to seek an answer to the question, What is it about truth that makes this relation hold?, the answer to which will involve telling us more about truth. In fact, it is usually thought that the answer to this question will involve providing a theory of truth that accounts for the peculiar explanatory role of that notion.

The claim that truth conditions have an explanatory role to play, or, more precisely, the claim that the fact that an utterance or belief has the truth condition that it has plays an explanatory role, can be explained in a similar fashion: if the proposition that Allison's belief is true iff  $p$  bears an explanation relation to some other proposition, then the fact that Allison's belief has the truth condition that it does has a role to play in the explanation of (e.g.) some fact about Allison's behavior. If truth conditions have an explanatory role to play, then it is not unreasonable for the substantivist to wonder about how they get assigned, and it is not unreasonable for her to seek a theory that would explain this phenomenon.

Although our discussion has been pretty general, I hope it is a bit more clear what it means to say that truth or truth conditions play an explanatory role. To say that they play an explanatory role is to say that the proposition that  $S$ 's belief is true, or that  $S$ 's belief has a particular truth condition, bears a particular relation—the explanation

---

<sup>132</sup> The discussion sometimes is couched in model-theoretic terms, i.e., in terms of truth or falsity on an interpretation. Hence, the claim would be that there is an explanatory relation between a belief's being true on the standard interpretation and something else.

relation (sometimes understood as the relation *is a cause of*, or *is a partial cause of*)—to something else, for example, the proposition that S succeeds at some task or other.<sup>133</sup>

### 3.3 Truth, Explanation and Deflationism

Deflationists have been very concerned to uncover the explanatory role of truth, arguing that if no explanatory role can be uncovered, then we have no reason to expect philosophers to unearth a substantial theory of the nature of truth. Inflationists—here I think of, among others, Alvin Goldman (1989) and William Alston (1996)—have been impatient with the deflationists, wondering (in print) why it is so important to uncover an explanatory role for truth to play.<sup>134</sup> From what was discussed above, it should be obvious why it is important to the deflationist that truth not have a substantial explanatory role to play: if truth (or truth conditions) has an explanatory role to play, then the deflationist, *qua* deflationist, is committed to semanticalism, the view that truth is a primitive indefinable, explanatory, and maybe even causally efficacious, property. The deflationist is a deflationist, in part, because she does not believe that truth has an

---

<sup>133</sup> Just to show that this way of characterizing things is not trivial, we should contrast it with an alternative “success” explanation. Rather than claiming that the proposition that S’s belief is true bears the explanation relation to the proposition that S succeeds, one who denies that truth plays an explanatory role in success explanations might argue that, in fact, it is not the proposition that S’s belief is true that bears the explanation relation to the proposition that S succeeds, but, instead, it is the proposition expressed by S’s belief that bears the explanation relation to the proposition that S succeeds. One who held this view would, of course, deny that truth plays an explanatory role in so-called “success” explanations. Cf. Leeds, 1995, for a development of this line of thought.

<sup>134</sup> I should note that Michael Devitt is a substantivist about truth who appears to be very exercised about the explanatory role of truth, attempting, as he has over the years, to uncover an important role for it to play in one or another of our explanatory enterprises. He has now isolated an explanatory role for truth to play that, as far as I can tell, begs the question against the deflationist. Per Devitt (1996, 1998), truth is important to the semantic project, the project of explaining meaning. Although I do not think that he is incorrect, I cannot see that this claim would impress the deflationists, since, as we will show, in order to accept a thorough-going deflationist position on truth, one must adopt a deflationary account of meaning.

important explanatory role to play;<sup>135</sup> if it ended up that truth was an explanatory or causally efficacious property, most (all?) deflationists would be committed to finding a reductive analysis of it.

I assume that the reason that the deflationist (*qua* deflationist) is committed to semanticalism if truth has an explanatory role to play is clear, but if it is not, here it is: on the deflationist's view, if there is a property of truth, it is a primitive, indefinable property, axiomatized (or implicitly defined) by the instances of the relevant schemata mentioned previously. If truth has an explanatory role to play then truth is a primitive, indefinable, and explanatory property; like electromagnetic phenomena, it is a basic feature of the universe. It is a refusal to accept the claim that, to paraphrase Fodor, when the physicists complete the catalogue they've been compiling of ultimate and irreducible (explanatory) properties of things, next to spin, charm, and charge, they will have to make room for truth and reference.

That is clear enough, but it is not the only reason the deflationist is concerned with the explanatory role of truth. The other reason, to which we will turn now, is of more general interest: it is why everyone should be concerned with the explanatory role of truth, as well as of any other property one cares to think about.

Those concerned with solving the age-old philosophical puzzle about the nature of truth—deflationists and substantivists alike—owe a debt of gratitude to the deflationists. The deflationists, motivated largely by Field (1972), Leeds (1978), Putnam (1978), and in more recent times, Horwich (1990) and Field (1986) (though I suspect that

---

<sup>135</sup> This general claim would be hard—and pointless—to support, since people are deflationists for lots of reasons. I suspect that the primary reason for adopting deflationism is skepticism about the possibility of carrying out the reductive programs that have been developed by the substantivists.

the debt goes back further, e.g., to the views of Quine and, most likely, the positivists), have argued that we should expect substantive, informative analyses of notions like truth, meaning, etc., only to the extent that we can establish the need for such analyses, and that, once established, the need will play an important role in determining the shape that an adequate analysis will have to take.

So, for example, these philosophers reject the claim, were anyone to make it, that our twenty-five hundred year long preoccupation with truth as a central metaphysical notion provides us with a reason to expect that it is amenable to an elucidation or reductive analysis. In order to show that truth is the sort of property for which we should expect an elucidation, we must show that there is some fact or set of facts connected in one way or another with truth that will go unexplained unless we provide such an elucidation, which fact or set of facts will determine what sort of elucidation is needed.

This, then, is one way in which the deflationists have motivated the need for an account of the explanatory role of the notion of truth: if you want to know what truth *is*, you should first figure out what truth *does*, for, having done that, you will find that truth is whatever performs the function that you have uncovered.<sup>136</sup> An addition that is sometimes made, but that is not essential to the motivation, is that our analysis of truth, like our analysis of the concepts of folk psychology, is an *explication*, where, in the provision of an explication, we do not attempt to determine the nature of an ordinary notion; instead, armed with an account of the legitimate function of (e.g.) truth, we introduce a new, restricted notion that is designed to do the theoretical work, over a

---

<sup>136</sup> This way of putting it owes a huge debt to the Ramsey-Carnap-Lewis style of providing theoretical definitions, though I do not believe that a commitment to this motivation for seeking an explanatory role for truth commits one to the sort of descriptivist account of reference that emerges from that picture.

limited domain, of our ordinary notion. If this is the goal, an account of the explanatory role of truth is required.<sup>137</sup>

I believe that numerous substantivists have balked at the notion that we should be overly concerned with the explanatory role of truth because they do not conceive of philosophy as being in the business of developing new and improved revisionist versions of our old workaday notions. It is not my interest to adjudicate this debate, which I suspect has played a large role in the history of twentieth-century philosophy, not to mention one difference between Wittgenstein, early and late. Instead, I want to point out two things. First, unless we have a sense of what truth is supposed to do, it is hard to see how we could evaluate the correctness of a candidate account of it (except maybe with respect to certain general features, such as consistency). Second, as I believe the deflationists see it (though they are not always explicit on this point), we should not assume that all notions are proper candidates for reductive analysis. We should only provide a reductive analysis if we have reason to expect one, and we only have reason to expect one if the explanatory role of the notion under consideration calls for it. Given that, we need some reason to assume that truth is a likely recipient. However, whether or not a notion should be reduced depends on its function, on what it does. Hence, if we are to determine whether or not truth is a candidate for reductive analysis, we must determine what it does, which is to say: we must determine its functional or explanatory

---

<sup>137</sup> Hartry Field has clearly embraced this motivation. In fact, it is because of this motivation that he advocates *methodological deflationism*, a view according to which we should assume that truth is deflationary and add on to our account only when, and insofar as, we must.

role. For these reasons, it seems that everyone—deflationist and substantivist alike—should accept the need to uncover the explanatory role of truth.<sup>138</sup>

So, the question that everyone concerned with truth must face is: what are the relevant facts about truth, and what must truth be like such that those facts can be explained?

The motivation for adopting deflationism comes in precisely at this point: We assume, following the deflationist, that we have a motivation for adopting a more-than-deflationary account of truth—that is, one on which the assignment of truth conditions to (e.g.) utterances is explained by reference to a theory of meaning, or of reference—only if we have discovered an explanatory role for truth to play that requires a more-than-deflationary notion of truth. We assume, with the deflationist, that the explanatory role of the property of truth is exhausted by that of our concept of truth—in effect, the truth predicate, ‘is true’—and that the function of our concept of truth is to enable us to capture generalizations that we could not capture using other familiar devices for forming non-semantic generalizations, such as the apparatus of objectual quantification.<sup>139</sup>

---

<sup>138</sup> I am not going to engage in a debate over the nature of explanation, though I realize that engaging in such a debate might be useful. I believe that the notion of explanation employed by some deflationists (e.g., Horwich) is rather strong. In the present case, since we are talking about the logical role of the truth predicate, the notion of explanation that is invoked by the deflationist involves a deductive relationship between explanans and explananda. I believe that the notion of explanation that is employed is very close to the notion of analysis that is employed in philosophical analyses (e.g., of truth, goodness, etc.). It is an a priori, philosophical notion of explanation that is far removed from the more pragmatic notion favored by some.

<sup>139</sup> This should be qualified a bit. We cannot capture these generalizations using the familiar devices *that are available to us in English*. Of course the truth predicate is only one of a number of devices that we could introduce to capture these generalizations. Please note that it is because of this that I balked at the explanations of truth in terms of substitutional quantification, propositional quantification, or something of that nature, that have been offered by deflationists: of course we can explain what truth is if we had these devices, but that is just because these devices would serve the same role as truth does. It was because of this that I took it as illegitimate to explain truth in terms of substitutional quantification, propositional quantification, etc.

We note that we do not require a more-than-deflationary account of the property of truth in order to explain the logical behavior of the truth predicate, so we conclude that if the logical function of the truth predicate exhausts the explanatory role of the property of truth, and we are assuming that it does, then we have no reason to expect a more-than-deflationary account of that notion. Hence, *modulo* the assumption about the importance of an explanatory role for the prospects of the development of a substantive account, the deflationist can establish that if truth has no explanatory role beyond the logical role of the truth predicate, then we have no reason to expect a reductive analysis of truth.

The deflationist actually has a stronger argument against the need to reduce, one that, if successful, provides us with a reason for thinking that no reductive strategy will be successful. The argument assumes—what must, of course, be shown—that the explanatory role of truth is exhausted by the logical role of the truth predicate, which role, as mentioned, is to serve as a device for formulating generalizations that could not be formulated, by us, in the language, as it is. If this is the function of truth, then, as is accepted by substantivists and deflationists alike, facts articulated using that concept could not be deduced from ordinary, non-semantic facts, from which it follows that they could not be explained in terms of them, and if they could not be explained in terms of them, then there could be no explanatory gain in supposing truth to have a substantive, underlying nature. It follows from this that we have no reason for believing that truth has a substantive underlying nature.<sup>140,141</sup>

---

<sup>140</sup> This argument was inspired by an argument made by Horwich (1995) about reference, though it is in important ways different from that argument (aside from the fact that my argument deals with truth and his deals with reference, of course).

<sup>141</sup> Of course one might balk at this claim, thinking that it relies on a very strong notion of explanation, but I think that is to miss the point. The claim does not turn on the acceptance of a very strong notion of

I believe that the conclusion that the deflationist should arrive at is even stronger than this, however: if (a) the facts about truth that need explaining are exhausted by the explanatory role of truth and if (b) the explanatory role of truth is exhausted by the logical role of the truth predicate, and if (c) claims formulated using the truth predicate cannot be formulated in non-semantic terms, then (d) we have every reason to expect that no analysis of truth *could* be forthcoming.<sup>142</sup>

If we are correct about the importance of establishing the existence (or non-existence) of an explanatory role, then questions of the explanatory liaisons into which truth enters are not of merely passing interest: they must be answered before we can even begin to think about what sort of a thing truth is.<sup>143</sup>

### 3.4 What is an Explanatory Liaison?

I think that the deflationist has been too quick to interpret the injunction to consider the facts about truth that need explaining as the injunction to consider the explanatory role of truth, i.e., what truth explains. Arguing from the putative fact that the explanatory role of truth can be wholly captured by a deflationary notion—one characterized by reference only to instances of the disquotational schema, together with

---

explanation. Rather, it turns on the fact that the role truth is playing is a logical role and that, as such, if we can express the facts articulated with the concept of truth (i.e., the truth predicate) without employing the apparatus of truth, we should be able to deduce our 'true'-containing generalization from our non-semantic generalization.

<sup>142</sup> Another way of arguing for the same point: in some circumstances, no reason to believe that p is not reason to believe not-p. I do not believe that this is one of those cases. In this case, no reason to expect a substantive account of the nature of truth is reason to believe that there will be forthcoming no substantive account of the nature of truth.

<sup>143</sup> Although I assume that it goes without saying, I in no way take this last section to have established or even made plausible the claim that the explanatory role of truth is exhausted by the logical function of the truth predicate. Indeed, we did not even consider the truth or falsity of this claim. In the next section, I will consider the question of whether or not truth is a causal property, and I will try to provide some reason for thinking that it is not. If I am right then one road to substantivism is blocked. As we will see, it is not the only road, and it is not the road that the substantivist should have been on, anyway.

uncontroversial facts and principles—the deflationist concludes that the conditions for expecting a substantive analysis of truth have not been met. This would be the right way to argue if by ‘facts about truth’ what was meant was *facts explained in terms of truth*, but this is not what is meant, or, at any rate, there is good reason to believe that this is not all that is meant.

I introduce the term *explanatory liaison* to cover both explanations *of* and explanations *in terms of* truth. So, one sort of explanatory liaison—the one that the deflationist points to— involves the explanatory role of truth, i.e., what is explained *in terms of* truth. Here, the deflationist’s answer is obvious: the explanatory role of truth is exhausted by the logical function of the truth predicate. The other side of the explanatory coin—the other sort of explanatory liaison—involves explanations of facts *about* truth, and the question that it raises is: Can we explain all of the facts about truth on the hypothesis that our theory of truth is as the deflationist claims?

In the argument given in the last section—(a) – (d)—the importance of premise (a) was to allow us to eliminate from our consideration this second sort of explanatory liaison. Hence, the argument turned on the un-argued for assumption that the claim that a deflationary theory of truth was explanatorily adequate amounted to the claim that it was adequate to capture the explanatory role of truth. Given the second explanatory liaison, we can see that this is an assumption that cannot be made. As we will see, there is reason to think that there are other facts about truth that need explaining, and that when we attempt to explain these other facts, our deflationist will be in trouble.

As I have characterized it, then, questions of the explanatory liaisons of truth can be of two sorts:

- Questions about the explanatory role of truth; or
- Questions about the explanation of the facts about truth.

Arguments about the explanatory role of truth, and about whether or not it supports a more-than-deflationist account of that notion, do not show that there could be no explanatory gain in supposing truth to be analyzable in other terms, i.e., do not show that a deflationary theory of truth is explanatorily adequate. If this is correct then our interest in assessing the adequacy of a deflationary theory of truth extends beyond the ordinary considerations of the explanatory role of truth. We must consider whether all of the facts about truth that need explaining can be explained on our deflationist's characterization of the nature of truth, which means that we must first determine what all of those facts are.<sup>144</sup>

### 3.4.1 Explanatory Liaisons of Type I: Explaining Behavior

Of course if we could show that truth does have a significant explanatory role to play, we would succeed in undermining the deflationist's position. I think that showing that truth has the sort of explanatory role that would support the development of a substantive account of truth is considerably harder than some philosophers have thought, and I think that it might prove useful if we ran through the reasons for this. In order to do so, I will consider a recent challenge to a deflationist view of truth that has been made by John Gibbons (1999). Although I will be arguing against Gibbons' particular attack on deflationism, the points that I adduce are quite general, and they can be used to assess the

---

<sup>144</sup> I assume we know what the deflationist says about the nature of truth, namely, that it does not have one, or, at any rate, that it does not have one that cannot be adequately characterized by the collection of all instances of (DS)

merits of any claim to the effect that truth has a significant role in the explanation of behavior.<sup>145</sup>

For the purposes of our discussion, it is not necessary for us to lean too heavily on any one version of the deflationist position. We can employ the broadly Horwichian view, which is the view, roughly, that all of the facts about truth that need explaining can be explained by reference to the various disquotational schemata, together with unproblematic principles and premises such as (EXP), i.e., the view that involved the propositions expressed by instances of the disquotational schema, together with a deflationary view of what it is for a sentence, belief, or utterance to express a proposition. The employment of propositions was, as previously noted, immaterial, but I will keep it as part of the view, because (i) I think that it can be independently motivated and (ii) I believe that it can be understood in terms that render it relatively harmless.<sup>146</sup>

### 3.4.2 Introduction to Gibbons' Argument

Gibbons' attempt to undermine deflationism—i.e., by pointing to the explanatory role of truth—is a good example of the sort of argument that has been given by other substantivists (e.g., Devitt (1984, 1996, 1997), and Alston (1996)). To my knowledge the deflationists have not yet addressed this specific sort of challenge, so there is independent interest in the argument I will give. Furthermore, as I hope to show, the issues that I raise might help to clarify the sorts of considerations that must be brought to bear on an

---

<sup>145</sup> More specifically, the argument that I give generalizes in very obvious ways to other alleged claims about the role of truth in the explanation of behavior (and successful behavior). I believe that the divide-and-conquer strategy that I advocate can be used by the deflationist to defend against other substantivist challenges as well.

<sup>146</sup> I should note that it might also prove necessary if we are to explain exactly what Gibbons means by a 'claim'. On one reading of that notion, a claim, like a statement, is a proposition.

assessment of any challenge to the explanatory adequacy of the deflationary theory of truth. Furthermore, they will show—what I believe to be the case—that it is tougher to slay the deflationist dragon than Gibbons and others appear to recognize.

The conclusion of Gibbons' argument is that truth is a causally relevant property. When this is combined with his claim that, on the deflationist's view, truth could not be a causally relevant property, the result is that truth is not as the deflationist would have us believe. Gibbons argues that the deflationist will attempt to resist his conclusion, i.e., by locating the relevant causal powers elsewhere—specifically, in the individual beliefs held by thinkers, together with the corresponding facts. He admits that this strategy is effective in some cases, but attempts to show that it is insufficient to handle all of the relevant cases. When the deflationist alters her account so as to capture all of the relevant cases, though she may successfully accommodate all of the cases without invoking the concept of truth, i.e., the truth predicate, she nevertheless reintroduces the property of truth and so undermines her own position. At any rate, this is what Gibbons attempts to show.

I think that Gibbons' argument rests on a mistaken view of what the deflationist can and cannot accept, and, although it is not my interest to defend deflationism against all comers, I think that it is important to point out where it is and where it is not vulnerable to criticism. In what follows, I will briefly recount Gibbons' argument, and will show why it poses no threat to the deflationist.

### **3.4.3 Gibbons' Argument: Truth in Action**

Gibbons characterization of the deflationary conception of truth is largely in accord with the version of deflationism that we are assuming, though it goes further than I believe any deflationist would or should be willing to go. According to Gibbons' deflationist (1999, p.4)

[the deflationary schema] is all there is to our notion of truth. There is, according to the deflationist, no interesting and important feature shared by the claim that snow is white and that money doesn't grow on trees. Similarly, there is no common feature shared by the claim that its true that snow is white and its true that money doesn't grow on trees.

If Gibbons means to say that the claims (propositions?), e.g., that snow is white, that money doesn't grow on trees, etc., share no interesting features, he is wrong. It should be clear that even the deflationist will claim that they share some interesting features. To take one example, Field would say that, assuming claims to be meaningful sentences, what's interestingly common to all of the aforementioned claims is that they are all sentences of my idiolect, and, therefore, all proper substituends for 'p' in the disquotational schema. Taking claims to be propositions, Horwich might say that these are all in the extension of 'true', which is a feature that they share, and one that is, *prima facie*, rather interesting, even to the deflationist.

What Gibbons should have said, and what the deflationist does accept is that there is no interesting or important property common to the aforementioned claims in virtue of which it is appropriate to ascribe truth to them. In other words, though all of the aforementioned "claims" are true (or, per Field, true-as-I-understand-them), it does not follow that there is some common feature,  $\phi$ , such that each is true *because* it is  $\phi$  (or

because it bears  $\phi$  to the world, a linguistic community, or what have you). Hence, contrary Gibbons, the deflationist does not deny that truth claims might share any interesting or important common features; rather, the deflationist denies that there are interesting or important common features of the claims to which truth applies in virtue of which they are, or are taken to be, true.

Having straightened out that matter, we return to Gibbons. Gibbons argues that truth is an interesting and important property, and that what makes it interesting or important is that it does causal work. What causal work? Gibbons considers “our ordinary way of thinking and talking about human action.” (1999, P. 1) He imagines Marcia and Greg, both of whom stop at red lights because they want to avoid getting traffic tickets. On one way of thinking and talking—Gibbons’ preferred way—Greg and Marcia act in the same way and do so for the same reason, though in many respects their acts are very different. For example, Marcia is on a Harley Davidson, whereas Greg is tooling around in the family car; Marcia stops by squeezing a hand brake, whereas Greg does so by applying pressure to the brake pedal; etc. These events, though they can be type identified in such a way that they count as the same, can be typed in such a way that they count as different, too, as Gibbons suggests.

Without going into too much detail, we should note that Gibbons (1999, p. 3)<sup>147</sup> points out that

---

<sup>147</sup> It probably bears pointing out that this is one place where Gibbons appears to be a bit confused about the nature of deflationism. If truth conditions are causally relevant then, presumably, we need to explain how sentences, utterances, or what have you, get them, and if we can explain how sentences get them, we are well on our way towards the possession of a bona fide inflationist theory of truth. How? Suppose that we explain how an utterance of ‘New York is populous’ gets its truth condition—that New York is populous: we claim that for it to have the truth condition that it has is for it to have some property, P. In that case, we could say that at least one way for an utterance to be true is for it to have P and for New York to be populous. The moral: we can always say what it is for a putative truth bearer to be true if we can

if we are concerned with systematic or lawlike connections between mental states and behavioral events, the relevant properties of the effects are the intentional behavioral properties. Now I claim that the truth value, not just the truth conditions or contents of your beliefs is causally relevant. Truth is a causal power.

Why is truth a causally relevant property, according to Gibbons? Marcia wants to stop at the red light and believes that by squeezing this lever, she will stop. If her belief is true, she will stop. If the belief had been false, she wouldn't have stopped.<sup>148</sup> Per Gibbons, this suggests that truth is causally relevant to ordinary end result types like *stopping at red lights*, as well as to their intentional counterparts.

Enter the deflationist. Gibbons recognizes that the deflationist need not claim that the fact that Marcia's belief is true plays any causal work. Put generically, the deflationist can explain Marcia's stopping by pointing to the fact that, for some p, Marcia believed that p and p. More specifically, though of course not completely, the deflationist can explain Marcia's stopping as follows:

(50) Marcia wanted to stop at a red light.

(51) Marcia believed that if she squeezed the lever, she would stop.

(52) If Marcia squeezed the lever, she would stop.

What is relevant about the explanation that involves (50) - (52) and statements like those is that it will never make room for the introduction of further facts that regard truth.

---

explain how it comes to be that the (putative) truth bearers get their truth conditions. Hence, the deflationist about truth must also be deflationist about the relation *has the truth condition*.

<sup>148</sup> Gibbons does not consider issues regarding causal preemption, and I will not either. Nevertheless, we should note that it is not necessarily true that if her belief had been false, her motorcycle wouldn't have stopped. For example, her motorcycle might have run out of gas.

Hence, if the deflationist can always use such particular statements in her explanation of the causes of behavioral events, she can undercut the claim that truth is a causally relevant property. None of this is news to Gibbons, of course.

As Gibbons sets things up, there is a clash: he maintains that in order to explain behavioral events like the aforementioned, we have to invoke the truth of (e.g.) Marcia's belief. The deflationist disagrees. She denies that, in our explanations of events like Marcia's stopping at a red light, we ever have to introduce a causally-relevant property of truth. In order to adjudicate, Gibbons returns to Greg and his family's station wagon. Like Marcia, Greg stopped at the red light; also like Marcia, he did so to avoid a traffic ticket. As Gibbons sees it, and I will not object, the causes in both cases involved the desire to avoid traffic tickets; the effects in both cases were of the type *intentionally stopping at a red light*; and, finally, another partial explanation of their stopping might involve pointing out that both Greg and Marcia have true beliefs about how to stop.

This last aspect of the cause of their stopping is crucial to Gibbons' argument. Gibbons' focus is the sort of case in which we want to provide the same explanation for both behavioral events. He notes that if we are trying to set out the similarities in the causes of their behaviors, we cannot advert to particular beliefs shared by Marcia and Greg about how to stop, for they have different beliefs about how to stop, since, after all, one of them wants to stop a Harley, while the other wants to stop the family station wagon. Hence, he argues that, in our explanation of the two behavioral events, we will have to advert to the fact that both Greg and Marcia have true beliefs about how to stop—nothing more specific will do.

Again, as Gibbons correctly notes, the deflationist has a response. If we want to explain their behaviors (typed as *intentionally-stopping-at-a-red-light* behaviors), we can note the following:

(53) Both wanted to stop at a red light.

(54) Both satisfy the following: for some p about stopping at red lights, S believes that p and p.

According to Gibbons, (54) allows us to slip in that, in each case, there is some mind-world relation such that part of the explanation (cause) of the behavioral event (effect) to be explained is that the agent—Marcia or Greg—has a particular, though unspecified, property—the property of believing that p when p.

As the deflationist would describe it, this is the property of having a true belief, for the notion of having a true belief is cashed out in terms of something like (54): S has a true belief iff for some p, S believes that p and p. But now Gibbons (*Ibid.*, pp. 5 -6) turns the tables on the unsuspecting deflationist, saying

But that means that this notion, the notion of having a true belief, *is* playing a causal role. If, as I have claimed, truth's being a causal power is inconsistent with deflationism then this strategy is not available to the deflationist. The deflationist does not use the word "true" in picking out the causal powers, but does use the notion of truth.

Thus ends Gibbons' argument. Has he shown that truth is a causal power, thereby undermining deflationism? Well, as Gibbons claims, if truth plays a causal-explanatory role then it too must be explained, and it looks as if Gibbons has shown that the notion of a true belief does play a causal-explanatory role.

### 3.4.4 The Deflationist Response to Gibbons' Argument

I do not think that Gibbons has undermined deflationism, however, and I think that there are at least three considerations attention to which will bear this out. In what follows, I will list the considerations and then I will attempt to extract responses to them from Gibbons' paper. After that, I will attempt to generalize the results of our findings, to show where I think anti-deflationist arguments from the explanation of behavior falter.

First, if truth plays a causal role, as Gibbons claims, we should ask just what role truth is playing in Gibbons' case, and, in particular, whether or not it is a role that the deflationists can either recognize or accommodate.

Second, Gibbons claims that, in order to undermine deflationism, all that must be shown is that truth is an interesting or important property. This may be true, but if it is, then there must be—and there is—a reason for this. If the fact that truth is an interesting or important property is to undermine deflationism, it must be shown that truth is interesting and important *in a way that could not be explained by the deflationist*. To undermine deflationism, it is not enough to point to some reason for thinking truth to be interesting or important.

Third, if truth is an interesting and important property that plays a role that cannot be accommodated by the deflationist, then there should be some sort of account of truth that is suggested by the interesting or important role of truth that is uncovered. Hence, we should ask what sort of account of truth would yield the sort of notion that Gibbons claims we need.

Let's leave to the side the possibility of explaining Marcia and Greg's behavior without adverting to propositional knowledge at all, i.e., as a bit of *knowledge how*,

rather than as *knowledge that*, and grant that the sort of explanations Gibbons calls for must be given. Gibbons notes that if we want to explain both Greg's and Marcia's intentionally-stopping-at-a-red-light behaviors, then we will advert to the fact that they both satisfy the existential claim made in (54). But since (54) captures the deflationist's characterization of what it is for someone to have a true belief, the deflationist notion of truth is implicated in the explanation of Greg's and Marcia's behaviors. Hence, Gibbons argues that his example shows that the deflationist's notion of truth is causally relevant after all, and, therefore, that it has a role that cannot be accommodated by the deflationist, since, as we know, the deflationist is committed to the view that truth is not a causally relevant property.

What role is truth playing in Gibbons' example? Well, we could not explain both Greg's and Marcia's behaviors by pointing to some particular causally efficacious, mutually held beliefs about stopping at red lights, since, *ex hypothesi*, their particular beliefs are very different. What we need is a way of capturing the sense in which each succeeds because s/he has some particular beliefs, though the beliefs had by the one might not be had by the other. Hence, rather than (51) and (52) (and correlative propositions for Greg), what we need is a (infinite) disjunction, as in (55).

(55) Marcia believes that if she squeezes the lever, she will stop and if she squeezes the lever she will stop and Greg believes that if he squeezes the lever he will stop and if he squeezes the lever he will stop OR Marcia believes that if she applies pressure to the brake she will stop and if she applies pressure to the brake, she will stop and Greg believes that if he applies pressure to the break, he will stop and if he applies pressure to the break, he will stop OR Marcia believes that if she squeezes the lever, she will stop and if she squeezes the lever she will stop and Greg believes that if he applies

pressure to the break, he will stop and if he applies pressure to the break, he will stop, ad infinitum.

Of course we have another option; we can use something like (54) or (56).

(56) Both have true beliefs about stopping at red lights.

So, what Gibbons' argument purports to show is that, if we want to explain Greg's and Marcia's behaviors as both being caused by beliefs of the same type, we must type identify them *semantically*, i.e., not just as beliefs about stopping at red lights, but as *true* beliefs about stopping at red lights. If this is correct, then, referring back to our discussion at the end of chapter 2, it suggests that truth is functioning as a genuinely descriptive term, viz., as a device for categorizing and explaining on the basis of shared characteristics, which, if true, would undermine at least one very important deflationist motivation.

On second thought, this appearance might be illusory, for if we can avail ourselves of (54), (56), or (55), then it might be that truth is playing its familiar role as a device of generalization, the role that the deflationist claims provides its *raison d'être*. This should make us a bit suspicious about the causal role that truth is playing in Gibbons' example. More specifically, notice that, though we do invoke truth (or the deflationist's rendering of truth) in our explanation that covers both cases, it does not seem as if *truth* is doing any causal work.

As we would describe the case, and granting the causal efficacy of psychological attitudes, what caused the event of Marcia's stopping was the fact that she had the attitudes that she had and what caused the event of Greg's stopping was the fact that he had the particular attitudes that he had, together with the correlative facts about the

world. We cannot deny the utility of the truth predicate in expressing these facts, but it looks as if the predicate is especially useful in this case because it allows us to express them in one fell swoop, without suggesting that what Marcia believed and what Greg believed are, when type identified in terms of content, the same thing. The concept of truth—whether it is captured explicitly, as in (56), described, as in (54), or cashed out, as in (55)—is there to help us to do what we could not otherwise do, unless our language came equipped with quantificational machinery that it lacks, or unless we had the ability to run through an infinite disjunction like the one started above, which we do not. So, it seems that the role that truth is playing is the familiar role as a device of generalization.

We can bring this out in a slightly different way. When we restricted our attention to Marcia's case, we saw that, aside from her desire to stop at a red light, what explained Marcia's behavioral event was the fact that she believed that if she squeezed the lever she would stop and, as happens to be the case, if she squeezed the lever, she would stop. When we restrict our attention to Greg's case, we notice that the same sort of explanation could be given. In both cases, what explained the behavioral event was the fact that the agent—Greg or Marcia—had the particular belief that s/he had, together with the proposition that is expressed by the agent's belief.<sup>149</sup> In the case where we aim to explain both behavioral events, we still believe, and intend to communicate, that what explains each behavioral event is (i) the fact that the actor has the belief that s/he has, together with (ii) the proposition expressed by his/her belief. That is to say, we want to explain the behavioral events by saying that for each agent, Marcia and Greg, there are

---

<sup>149</sup> In this sentence, I wrote that it was the fact that the agent had the belief that s/he had, together with the *proposition* it expressed, that explains the behavioral event. 6 sentences ago, I wrote that it was the fact that s/he had the beliefs that s/he had, together with the *facts*, that explains it. This switch—from talk about

propositions,  $p$  and  $q$  which are about stopping at red lights (where, possibly,  $p \neq q$ ), such that Marcia believes that  $p$  and  $p$  and Greg believes that  $q$  and  $q$ . Alternatively, if we avail ourselves of what Dorothy Grover (1992) calls *philosopher's English*, we can say (54), or we can begin the infinite disjunction, (55). In any case, truth is playing its familiar role as a device of generalization. It is still the particular propositions, together with the belief in such, that explain the aforementioned behavioral events.

Next, we turn to the question of whether or not truth is an interesting or important property in a way that cannot be handled by the deflationist. Evidence that it is not is easy to find: In order to show that truth is an interesting or important property, Gibbons has relied on the deflationist's account of the "cash value" of 'true belief'. Call the notion of truth implicated therein—the notion found in (54)—'truth<sub>d</sub>', i.e., to indicate that this is the notion that the deflationist endorses, since we have shown that it is. If the claim that truth is an interesting or important property is to undermine deflationism, it must be shown that the notion of truth that is interesting or important is not truth<sub>d</sub>. The argument would be straightforward: truth is an interesting or important property; *ex hypothesi*, truth<sub>d</sub> is not an interesting or important property; hence  $\text{truth} \neq \text{truth}_d$ . The problem is that Gibbons has claimed that truth<sub>d</sub> is an interesting or important property. It may well be, but that is all the better for the deflationist.

Finally, we consider the sort of account of truth that seems to go with Gibbons' argument. Here, it might seem that the deflationist is in trouble, for, rather than (54), (55), or (56), something like (57) might seem appropriate.

(57) Both Marcia and Greg have beliefs about stopping at red lights that correspond with the way things are.

More to the point, it might seem that what unite (54) - (57) is that putative fact that they all capture the sense in which part of what explains Marcia and Greg's behavior is a certain sort of mind-world relation the obtaining of which, as Gibbons puts it, explains why "some, but not all, beliefs are true."

I do not think that we can deny that there is a sense in which it seems appropriate to explain Greg and Marcia's behavior by pointing to the fact that each has beliefs that represent the way the world is. However, I think that it would be a mistake to read too much into this. For one thing, everyone, deflationist and inflationist alike, can accept (57). The thought that the deflationist cannot endorse (57) involves a mistaken view of the difference between deflationism and substantivism akin to the mistake made by philosophers who claim that deflationists and substantivists accept different kinds of truth conditions. Of course deflationists do not countenance different *kinds* of truth conditions than are recognized by substantivists; rather, as previously mentioned, they accept different *explanations* for why it is that our sentences, utterances or beliefs have the truth conditions that they have.

There is a parallel in the present case. What matters isn't whether or not one accedes to (57), but is, instead, what account one gives of the relevant correspondence relation, for, on at least one reading, the reading endorsed by the deflationist, of course, (57) is platitudinous. The deflationist does not deny that we can engage in representation talk, after all; what she denies is that we can provide a naturalistically acceptable, or at least a non-trivial, elucidation of the notion of representation. Hence, she will deny—what Gibbons appears to assume—that there is a mind-to-world representation relation, *R*, such that a thinker has true beliefs when and only when she bears *R* to a chunk of the

world. Hence, I cannot see that the fact that we unify (54) - (56) in terms of something like (57) in any way threatens the deflationist.

What non-deflationary account of truth does Gibbons' argument seem to recommend? If we are right about the role that truth is playing in Gibbons' argument, and if we are right that he is merely showing us that there is a sense in which we might think that  $\text{truth}_d$  is important, then the account that is suggested by his argument is at least a notion like  $\text{truth}_d$ . But a deflationary account of truth yields  $\text{truth}_d$ . Hence, I cannot see that his argument suggests the need for anything beyond a deflationary notion of truth, in which case, even if he has succeeded in showing truth to have certain causal powers, which is a contention that we have given reason to question, it does not follow that he has shown a deflationist account of truth to fail to be up to the challenge of accounting for the explanatory role of truth. I conclude, therefore, that Gibbons' argument poses no threat whatsoever to the deflationist.

I believe that a careful consideration of the three questions noted above—viz.,

- Is the role that truth is being claimed to play one that the deflationists cannot recognize or accommodate?
- If truth is an interesting or important property, is it interesting and important in a way that could not be explained by the deflationist?
- If truth is an interesting and important property that plays a role that cannot be accommodated by the deflationist, what should truth be like such that it can play that interesting or important role?—

will bear out the fact that it is considerably harder to argue against the deflationist than some philosophers have thought. I fear that substantivists have not given these questions

enough thought, the result being that many of the arguments against the deflationary theory of truth have not moved the deflationists to alter their position at all. To conclude this section, then, I will attempt to draw out a more general moral regarding the role of truth talk in psychological explanation.

### 3.4.5 A General Moral Regarding Psychological Explanation and Truth

Having gone through one argument for the claim that truth has an explanatory role (in fact, a causal-explanatory role) that cannot be accounted for by the deflationists, I will not proceed to the myriad variations on this theme that a substantivist might put forward. However, I would like to close with one final thought, a thought related to our solution to the problem posed by Gibbons. It is often argued—Cf., especially, Field, 1986, though I fear that the argument there is confusingly laid out<sup>150</sup>—that the deflationist cannot accommodate the fact that, in some sense, we rational agents succeed in our intelligent behavior *because* we have true beliefs.

The argument, distilled down to its essentials, is this: consider some action, A, such as the act of solving a crime. Sometimes we explain how it is that an agent, S, succeeded in performing A by citing lucky guesses; however, sometimes we explain how S succeeded in performing A by citing the fact that S has true beliefs about how to achieve A. We claim that S succeeded in performing A because she had true A-related action-guiding beliefs.<sup>151</sup>

---

<sup>150</sup> I found the argument confusing, and, to his tremendous credit, Field openly acknowledges it, i.e., both in conversation and in print.

<sup>151</sup> I am avoiding complications, by suppressing the relativization of the interpretation of the beliefs to a standard interpretation. One version of the substantivist's argument involves the claim not simply that S's beliefs are true, but that they are true relative to the standard interpretation. On that version of the argument, the question is: What makes the standard interpretation special? The deflationist claims that there is nothing that we can point to—no causal relation, for example—that singles out the standard

This claim is thought to cause problems for the deflationist for the following reason: if S succeeds *because* her A-related action-guiding beliefs are true, then we want to know what it is about the property of truth that makes it the case that beliefs that have it can play this very important causal-explanatory role.<sup>152</sup> That is, we want a theory of truth, or of the assignment of truth conditions to utterances and thoughts, that will provide us with a principled way of distinguishing true beliefs from all other beliefs, where what distinguishes the true beliefs from the others will provide us with the means necessary to explain how it is that they can play the special explanatory role that the substantivist alleges that they play.

I am not convinced by the so-called success arguments, and, though I will not endeavor to go through them all, I will cite what I think is a general problem that plagues every one that I have encountered thus far. I think that the claim, that S succeeds at performing A because her A-related action-guiding beliefs are true, is ambiguous.<sup>153</sup> It is useful to invoke propositions at this point, as we did when we were discussing a related issue having to do with Gibbons' discussion of Marcia and Greg. The ambiguity is between the claim that what explains S's success is the proposition that her A-related

---

interpretation. He claims further that there is no reason to expect the standard interpretation to have some unique feature in terms of which it could be distinguished from all other interpretations.

The inflationist points to the so-called success argument that I am sketching, in order to show that there is something that makes the standard interpretation special: it, and it alone, is the one that is relevant to S's success. The inflationist takes this as evidence that there is sufficient reason to expect there to be some naturalistically describable feature of our language use (or thought) in terms of which the standard interpretation can be distinguished from all others.

<sup>152</sup> I assume, for present purposes, that we think in a language of thought, as is both standard (as well as dispensable) in discussions of this issue. This is just so that we can talk about S believing sentences, which I find easier than the other options. I assume, also, that when we say that S succeeded because she believed true A-related action-guiding sentences, we intend 'because' in such a way that our claim supports (or is equivalent to) the claim that the fact that S believes true A-related action-guiding sentences explains A's success.

action-guiding beliefs are true, and the claim that what explains her success are the propositions expressed by her A-related action-guiding beliefs, together with the fact that she has the A-related action-guiding beliefs that she does.

The substantivist will resolve the ambiguity in favor of the former reading, which explains why her interest in S's success will lead her to try to find out all that she can about truth; the deflationist will resolve the ambiguity in favor of the latter reading, which explains why her interest in S's success will lead her to find out all that he can about sleuthing (together with how S came to hold the A-related action-guiding beliefs that she holds, perhaps). Either philosopher could be described as interested in explaining the putative fact that S succeeds at performing A *because* she has true A-related action-guiding beliefs, what distinguishes them is the fact that the deflationist divides explananda in a way that the substantivist does not.<sup>154</sup>

I think that the deflationist would be right to push the second reading of the claim that S succeeded at A because her A-related action-guiding beliefs are true. More importantly, however, I think that the substantivist needs to provide a reason for believing that the second reading is insufficient, and, to date (and as far as I know), no inflationist has said anything to suggest that the deflationist's way of dividing explananda is unacceptable. Hence, I am doubtful that the sorts of arguments pulled from considerations of psychological explanation (e.g., by Devitt, 1984, Field, 1986, and

---

<sup>153</sup> Stephen Leeds (1995, p. 22) makes a similar point, although he is concerned with a model-theoretic version of the argument. I follow Leeds in talking about this in terms of propositions, and in taking explanation to be a relation between propositions.

<sup>154</sup> It is worth pointing out that the way in which the deflationist will divide explananda is exactly the way in which Gibbons says that she will, i.e., in terms of (51) and (52).

others) are going to undermine the deflationist's position on the explanatory role of truth.<sup>155</sup>

All of that said, I am not yet prepared to grant that the deflationist can accommodate all of truth's explanatory liaisons. That is because I am not convinced that a deflationary account of truth is sufficient to explain all of the facts about truth that need explaining. It is to this issue that we turn next.

#### **Chapter 4: Truth and Meaning**

Ever since Frege, and with only a few noted exceptions,<sup>156</sup> philosophers have thought that there was an intimate connection between the meaning of a sentence and its truth condition. On some accounts, e.g., Davidsonian accounts, the connection is straightforward: since Davidson (e.g., 1967) takes a (Tarskian) definition of truth for a language, L, to give the meaning of every sentence of L, he takes a theory of truth for the language to serve as a theory of meaning for that language, in which case the meaning of a sentence just is its truth condition. On other accounts, the relationship is a bit less

---

<sup>155</sup> There are other sorts of arguments that substantivists might use to undermine the deflationist's view of the explanatory role of truth, but I will not pursue them here. If what I have said is correct, then arguments from psychological explanation will not undermine the deflationist's position on the explanatory role of truth.

<sup>156</sup> One of the noted exceptions is Quine, who despaired of talk of meaning. Another of the noted exceptions, of course, is Katz, who believes that, though the meaning of a sentence (token) encodes a necessary condition on its truth, it does not encode a sufficient condition. Hence, Katz does not see meaning as being part of the theory of reference and truth; rather, he sees it as the subject of an autonomous theory that is related to the theory of reference and truth *via* referential correlates, which function as inter-theoretical bridge laws. The *job* of a theory of meaning—a theory of sense—, according to Katz, is to explain sense relations and properties, properties such as synonymy, analyticity, redundancy, and the like.

Recalling our discussion in chapter 0.1 about the way in which the deflationist might divide explananda, we can see that the questions that Katz's sense theory is intended to answer are the *internal* questions. Katz is not trying to answer the question of where sense come from; of how expressions get the senses that they have; etc. I believe that there has been much confusion about Katz's view stemming from the fact that the internal/external distinction has not been clearly recognized, for, in one sense, he is simply not trying to answer the question that (e.g.) Grice was trying to answer, about how meaning could arise. As he would put it, the difference is a difference of starting point, the top-down vs. The bottom-up approach.

snug, though it is close, nevertheless, for on these accounts—Fregean and Russellian accounts, in particular—the meaning of a sentence encodes, or in some way captures, that sentence’s truth condition. The relationship between meaning and truth conditions on this picture is relatively straightforward, though it is rarely made explicit. As I understand it, if the meaning of the sentence encodes that sentence’s truth condition—as it would, for example, if the meaning of a sentence was the Fregean or Russellian proposition it expressed—then the fact that a sentence has the truth condition that it has explains its having the meaning that it has.

Put a slightly different way, as Horwich (1998) notes, if meaning encodes truth conditions, then we are committed to the truth of, what he calls, *meaning-to-truth conditionals*, i.e., (M-T), which capture, or purport to capture the (M-T)  $s$  means that  $p \rightarrow s$  is true iff  $p$  relationship between meaning and truth. According to Horwich (1997, p. 100, 1998, pp. 27-30), these meaning-to-truth conditionals are very important. As he sees it, the main difficulty in resolving issues about meaning has been the need to explain *aboutness*, or *representation*, and the aboutness or representation—in a word, the *intentionality*—of language is captured in the meaning-to-truth conditionals. If Horwich is right, and we will assume that he is,<sup>157</sup> then resolution of the main difficulty regarding meaning requires explaining the instances of (M-T). Furthermore, if, as Horwich (N.B., and Soames, 1995, 1997, 1998) initially assumes, we are committed to the truth of all of the

---

<sup>157</sup> We might revise his claim a bit, holding that explaining the intentionality of language is one of the central issues in the philosophy of semantics. I make this emendation in light of the fact that on some accounts, e.g., Katz’s, the problem of intentionality is not, *per se*, a domestic problem for a theory of meaning, but is, instead, a foreign relations problem, and hence, is not the concern of a theorist of meaning.

instances of (M-T), then we have at least some reason to expect an explanation for why they hold, that is, an account of the relationship between the meaning of a sentence and its truth condition. We need an answer to the following question: what is it about meaning and truth that makes true all of the instances of (M-T)?

As I will show, Horwich's attempt to answer this question, and, more generally, to solve the problem of intentionality, so-called, fails. He cannot explain the connection between the meaning of a sentence and its truth; at least, he cannot do so in a way that is consistent with his deflationary theory of truth. As I will show, there is good reason to believe that this is not a problem unique to Horwich's account: no thorough-going deflationist<sup>158</sup> can explain the meaning-to-truth conditionals. The reason for this is that no through-going deflationist about truth can both accept the meaning-to-truth conditionals and avail herself of a more-than-deflationary account of meaning. Since no realist about meaning can comfortably deny the meaning-to-truth conditionals (or, at least, the fact that we accept them), it will follow that the thorough-going deflationist about truth must be a deflationist about meaning. This might not be news to all (though I suspect that it would be news to many, including Horwich), but it is important, for it increases the burden on the deflationist, who now must show both that she can provide a deflationary account of meaning and that there is proper motivation for so-doing. We will turn to this issue in the next chapter. First, however, we will turn to Horwich and the meaning-to-truth conditionals that he purports to explain.

---

except insofar as that theorist is concerned to relate the theory of meaning to other important theories, such as the theory of reference or truth.

<sup>158</sup> Just to remind the reader: a through-going deflationist adopts a deflationary view of truth for all putative types of truth bearer. Recall that the thorough-going deflationist is to be contrasted with the propositional deflationist.

There are two ways in which the deflationist might handle the meaning-to-truth conditionals. The first option, attempted by Horwich (1997, 1998), is to provide an alternative explanation for the instances of (M-T), one that makes use of a *use* theory of meaning, for example. We will consider Horwich's alternative in the next section. The other option is to accept the instances of (M-T), but to claim that they are, in some sense, *basic*. We will turn to this option, after having discussed Horwich's.

You might recall that we have distinguished between two senses in which a claim can be basic: it can be *conceptually* basic, meaning that we accept it in the absence of inference or argument, or it can be *explanatorily* fundamental, meaning that there is no possible explanation for why it holds. Horwich claims that the instances of (DS) are conceptually basic, and, as we saw, this was an important feature of his view, for it allowed him to claim that, in giving his deflationist account, he was explaining our ordinary conception of truth. In virtue of the fact that he denies that there is a substantive explanatory role for truth, he also took instances of (DS) to be explanatorily basic, which means that the fact that they hold does not admit of explanation.

There is a sense in which Field, too, takes the instances of (DS) to be conceptually basic.<sup>159</sup> Furthermore, there is a sense in which he would take them to be explanatorily fundamental, too, though for reasons very different from Horwich's.<sup>160</sup> I

---

<sup>159</sup> This might be a surprising claim, for Field is quite clear in his claim that we accept the instances of (DS) *because* of the cognitive equivalence of the left- and right-hand sides. However, you might recall that Field's notion of cognitive equivalence hardly deserved the name, for, as he understands it, to call two sentences cognitively equivalent, for a speaker, S, is for S's inferential procedures to license, fairly directly, the inference from one to the other and back again. In particular, there is not any interesting sense in which two sentences are cognitively equivalent because they share the same meaning. It follows from this that, per Field, we do not accept instances of (DS) because we accept any further, reductive premises, as we would on an account on which instances of (DS) were not conceptually basic.

<sup>160</sup> It is tricky to pin down Field on this issue. He tells us that the logical truth of the instance of (DS) follows from their cognitive equivalence, which means that it follows from the fact that our inferential

believe that Field would take them to be explanatorily fundamental in the sense that he did not think that we needed to provide any explanation for why they hold, or are true. However, I think that he believes this because he would deny that there is any interesting or important sense in which they actually do hold, or are true. We will discuss Field's view, after going through Horwich's.

Before we turn to the meaning-to-truth conditionals, we need brief introductions to the accounts of meaning of which the thorough-going deflationist might avail herself. For this, we turn to Horwich and Field.

Horwich's option is to provide an alternative to the sorts of reductive accounts of meaning favored by the substantivists. What distinguishes Horwich's account from the accounts favored by the substantivist is that he does not suppose that there is some relational, non-semantic property,  $P$ , in terms of which **meaning** can be defined, as in (P)

$$(P) (\forall x)(\exists y)(x \text{ means } y \leftrightarrow P(x,y)),$$

where  $x$  ranges over expressions of the language and  $y$  ranges over concepts.

Instead, Horwich accepts that there is a use-property,  $U$ , such that

$$(U) x \text{ means } y \leftrightarrow U(x),$$

where  $x$  and  $y$  are as above. Hence, Horwich's theory *is* reductive, though he does not attempt to reduce all meaning properties to some one, relational property,  $P$ . This theory

---

procedures sanction a (more or less) direct inference from either side of the biconditional to the other. This makes it sound as if instances of (DS) are conceptually basic. Of course, Field does not think that we can explain why our inferential procedures license the aforementioned, and, because of this, the instances of (DS) should be thought of more or less explanatorily basic. As I mentioned previously, I think that the most sensible thing for Field to say is that all he needs to explain is why we accept them, and that the question of whether or not they hold is sort of a don't care. In fact, I think he should claim that there is really no interesting sense in which they do hold, i.e., that he should adopt some form of nonfactualism about meaning *and* truth.

is an attempt to explain the *external* question about meaning, which, as Horwich puts it, is How does a mark on a page get the meaning that it has?

Horwich also thinks that his use-theory of meaning answers *internal* questions, such as questions about synonymy, antinomy, etc. (see especially Horwich, 1998, p. 49). Of course this is as expected. We would expect a particular answer to the external question to constrain the range of answers to the internal questions. It is important to note that Horwich believes that his use-theory of meaning, although it is reductive, does not provide the materials necessary to give a reductive analysis of truth, reference, or the 'true of' relation (predicate reference). It is because of this that he believes his use theory of meaning is compatible with his deflationary view of truth: we explain how it is that words and sentences get the meanings that they have, though not in a way that allows for the provision of an explicit definition of truth, reference, etc. As we will see, there is reason to think that Horwich's use theory of meaning does more than he thinks and that, in fact, it is incompatible with a deflationist view of truth.

Field opts for a deflationary attitude towards meaning; i.e., he denies that there are either objective or inter-personal meaning relations, and he deny that we can provide anything like a *theory* of meaning at all. We will discuss Field's account of meaning in more detail later, but, for now, it is important to note two things. First, the claim that Field's account of meaning is deflationary is a bit misleading. Field holds that, in some sense, meaning is *substantive*, for he adopts something like a conceptual role account of meaning, though he denies that we can make too much sense of the notion of sameness of conceptual roles across persons or times. Famously, conceptual roles do not determine

truth conditions; hence, Field does not hold that meaning determines truth conditions.<sup>161</sup> This leads to the second important feature of Field's deflationary view of meaning. Though it is not deflationary in the sense that he denies that our words have meaning (for us), it is deflationary in the sense that it denies truth or truth conditions any role in meaning. Hence, we might describe it not as a deflationary theory of meaning, but as a theory of meaning that is deflationary insofar as it denies that truth conditions—that clauses, as we will see—play a role in a theory of meaning. Field's view is not open to the criticism that we will level against Horwich's account, because Field does not offer anything approximating a reductive *theory* of meaning.<sup>162</sup> However, Field's view will raise other questions.

So, we see that there are (at least) two options available to the deflationist concerned with what account of meaning to adopt. The first option is to provide an alternative theory of meaning, one that (allegedly) does not interact with truth in the way that the traditional theories of meaning do. The other alternative is to provide no theory of meaning, i.e., to claim, in effect, that everything that needs explaining can be explained without resorting to anything like a theory. In what follows, we will evaluate these two options, first, by considering whether or not, in fact, the deflationist can avail

---

<sup>161</sup> I should note, briefly, that in his influential article "An Advertisement for a Semantics for Psychology" (1986), Ned Block argues that, though conceptual role does not determine reference, it determines which theory of reference is true (Block thought that the conceptual roles of the terms of our language determined that a causal theory of reference was true). Block was mistaken. His argument depended on the fact that he only considered there to be two accounts of reference available, a causal account and a description-theoretic account. Given only two options, conceptual role selects one by being incompatible with the other. What Block should have said is that conceptual role cannot determine reference, but it can determine which theory of reference is *false*. I take this to be the lesson of his (1986) discussion.

<sup>162</sup> I take it that that is why his paper on the subject is called "Deflationary *Views* of Meaning and Content." (emphasis added).

herself of an alternative theory of meaning and, second, by seeing whether or not she can avail herself of the no-theory option.

As we will see, there is reason to be pessimistic about both possibilities, which will lead us to accept the following conditional: If these two approaches, as developed, exhaust the options available to the deflationist, then, contrary to what the deflationist would have us believe, she cannot account for all of the facts about truth that need accounting for. There are two things to note about this conditional. First, if the antecedent is true and our arguments go through, we have shown that, attractive though it may be, generalized deflationism cannot be adopted. In light of this possibility, I will consider what about deflationism strikes me as obviously correct. Second, I will not argue for the truth of the antecedent. Even if there are alternative accounts that the deflationist about truth might co-opt,<sup>163</sup> there should still be much interest in the claim that these accounts are untenable. It is not my goal to show that no deflationist theory of meaning could possibly be developed, for I am not in a position to make that claim. That said, we turn to the deflationist's attempts to handle the meaning-related issues we have raised.

#### **4.1 Why Meaning Matters to the Deflationist: Jackson, Oppy and Smith**

One way to bring out the importance of meaning to Horwich's deflationary account of truth, and to thorough-going deflationist accounts, generally, is to consider a recent objection to deflationist accounts of truth that was made by Jackson, Oppy and Smith (1994). The objection turns on the question of which objects the deflationist takes as the primary bearers of truth. According to Jackson, Oppy and Smith, if the deflationist

takes propositions to be the primary bearers of truth, her view is trivial,<sup>164</sup> and if she takes sentences (or utterances) to be the primary bearers of truth, her view is false. A consideration of their argument, and of some of the available deflationist responses, will make clear the importance to the deflationist of meaning, and, in particular, of the development of an account of meaning.

In order to bring out the argument made by Jackson, et al, we should consider the claim, (49).

(49) *Snow is white* is true if and only if snow is white.

Jackson, Oppy and Smith ask what the deflationist takes the italicized sentence to pick out, a sentence or a proposition, and they take either choice to lead to trouble. First, they assume, with the deflationist (CF. Field, 1986, 1994a, 1994b; Horwich, 1990, 1993, 1995, 1997, 1998; and Leeds, 1974, 1978, 1995), that (49) should be read as making a necessary claim. However, if (49) is so read, then, if sentences are taken to be the primary bearers of truth, it is false—as they point out, the truth of the sentence ‘snow is white’ depends on more than just the whiteness of snow; it also depends on what the sentence *means*, which further requirement is not captured by (49). Suppose, then, that *snow is white* is taken to pick out a proposition—the proposition that snow is white. In that case, the theory is not false, for, as we understand the notion of a proposition, it is a necessary truth that the proposition that snow is white be true if and only if snow is white. However, now the accounts looks to be trivial, for the proposition that snow is

---

<sup>163</sup> For example, Schiffer’s Pleonastic view.

<sup>164</sup> In the final chapter of this dissertation, I will provide some reasons for being skeptical about this claim, though we will not consider it critically here.

white is *defined* as having the truth condition that snow be white, arguably, that is part of what makes it the proposition that snow is white.

As Jackson, Oppy and Smith see it, the deflationist has a dilemma: her view is either false or trivial. How should the deflationist resolve the dilemma? Assuming, for the moment, that they are right, our first inclination is to think that the deflationist should opt for the propositionalist option. One reason for this might be that we think that a trivial theory is better than a false one. A better reason is that we think that the claim that propositional truth is trivial is exactly the sort of thing that the deflationist would accept anyway, for, as we know, the deflationist believes that the most interesting thing about truth is how much less interesting it is than philosophers have thought.

Although this option sounds good at first, and although ultimately it might prove to be the better of the two options, the thorough-going deflationist cannot adopt it on the grounds that she is happy with the claim that truth is a trivial property. The reason is that, although the deflationist does accept that truth is, in some sense, a trivial property, she accepts this claim for reasons pertaining to the concept of truth. The triviality noted above follows from facts about how we are understanding propositions, which, as normally understood, are mind and language independent entities that have whatever truth conditions they have essentially.<sup>165</sup> Hence, although the claim that truth is trivial might be something to which the deflationist is attracted, if she is to argue for that claim, her argument should be based on things we believe about the concept of truth, rather than because of things we—substantivists and deflationists alike—believe about what it is to be a proposition.

There is a more serious issue, however. If the thorough-going deflationist takes propositions to be the primary bearers of truth, she could not possibly rest with the claim that propositions like (49) exhaust her theory of truth, for in so doing she would neglect the issue of meaning, which is to say, she would fail to develop an account of the relationship that a sentence bears to a proposition such that the proposition captures the meaning of the sentence. An account of propositional truth, if it is trivial, is trivial because it exchanges hard questions about what it takes for a sentence, belief, utterance, claim, etc., to be true (or false) for hard questions about what it takes for a sentence (etc.) to express the proposition that it does, which hard questions need to be answered by the thorough-going deflationist.

Now, our propositional deflationist might agree that it is important to develop an account of meaning, but she might hold that doing so is not her job—she is concerned with developing an account of truth, not with developing an account of everything. This response is insufficient, however, if our deflationist is a *metasemantic deflationist*. First, and most importantly, the standard accounts of meaning developed in the Fregean tradition are incompatible with deflationism about truth,<sup>166</sup> for they take meaning to determine semantic value. If our deflationist is to support her position on truth, she must show us that there is an account of meaning of which she can avail herself. Second, since the propositionalist will claim that utterances, sentences, beliefs, claims, and the like, are true or false depending on the truth or falsity of the proposition they express, if the deflationist is interested in explaining truth for all putative truth bearers, basic and non-basic, alike, then, having explained what it is for a proposition to be true, she has

---

<sup>165</sup> See Schiffer, 1996, for an excellent discussion of propositions.

barely begun her project. If she is a thorough-going deflationist, she adopts her deflationary view of propositional truth with an eye towards impugning the standard conception of a whole family of concepts, including reference, denotation, and presumably, meaning. For these reasons, a through-going deflationist, if she is to take propositions to be the primary bearers of truth, must develop a deflationary account of meaning.

Unimpressed with propositions, our deflationist might opt for a sententialist position, which is to say that she might take sentences (or, utterances) to be the primary bearers of truth and falsity. In that case, she must explain away the appearance that her view is obviously false (we will discuss this issue when we turn to Field). Of course, even if she takes this option, and even if she can stave off the charge of obvious falsity, she, too, must contend with the issue of meaning. Hartry Field offers a version of deflationism that, if successful, satisfies both of these desiderata. On Field's account, truth is ascribed to *interpreted* sentences, sentences that already have meanings. This certainly avoids the worry that the deflationist is ignoring meaning. It also explains why the putative falsity of (49), when the italicized sentence is taken to be a sentence, is only apparent. Of course, the sententialist is not claiming that the sentence token 'snow is white', understood as an uninterpreted string, is true if and only if snow is white; rather, she is claiming that the sentence, as interpreted, has that truth condition, and, in fact, that, interpreted as it is, it is necessary that it have that truth condition. As previously mentioned, in this instance, she is concerned with the issues that arise *within* the language—what we were calling the *internal questions*—rather than the philosophical

---

<sup>166</sup> They need not be. See, for example, Schiffer, 1999, "Pleonastic Fregeanism".

questions that can be asked from, so to speak, a God's eye perspective—what we were calling the *external* questions.

If the deflationist adopts this position, she takes on three heavy burdens. First, since truth attaches only to interpreted sentences, the tenability of her account turns on her ability to provide an adequate account of meaning. Second, she must provide an account that does not employ truth or truth conditions, as Field (1994a, p. 251) says, “under any description.” The reason is that if she does appeal to the concept of truth in her account of meaning, then, since her account of truth presupposes that truth-apt sentences are already interpreted or meaningful, it would be circular.

I imagine that, in reporting these two burdens, I am rehearsing what the deflationist takes to be old news. The third burden has not been explicitly noted by the deflationist, and there is good reason to believe that it has not been noted at all. The third burden involves the deflationist's provision of a non-truth conditional account of meaning. On that view, we cannot say what the meaning of a word, phrase or sentence is because there is a sense in which words, phrases and sentences do not have meanings. Of course they are *meaningful* to speakers that understand them, but that just means that they have conceptual roles in the speakers' idiolects. As we might put it, the deflationist about meaning denies that there are *objective* meaning facts, facts that could be stated by means of a 'that'-clause. The burden carried by such a position, stated very briefly (we will return to it later) is this: As for truth, so for meaning. The deflationist about truth has as a crucial premise in her argument a putative datum about the explanatory role of truth. This putative datum explains why we have a truth predicate in the language in such a way that the deflationist can claim that no reductive account of truth is called for.

We also have a meaning predicate in the language, and so it is reasonable to expect that a deflationist will have to provide a linguistic role for it, given that, when a speaker informs us (e.g.) that ‘snow’ means *snow*, contrary to what we might have thought, she is not ascribing to it some objective property.

The upshot of our discussion is that the dilemma posed by Jackson, Oppy and Smith, although it does not undermine either the propositionalist or the sententialist version of deflationism, raises exactly the questions that we will turn to next:

- Can the propositionalist—e.g., Horwich—explain in a suitably deflationary way, the meaning-to-truth conditions, N.B., which Horwich takes to capture the “representational power of language”?
- Can the sententialist—e.g., Field—provide us with a non-truth-conditional account of meaning, and can he explain away our intuitions about the contingency of sentential truth conditions (i.e., the intuition that the truth of the sentence ‘snow is white’ depends on more than just the color of snow)?

#### **4.2 Horwich’s Account, Briefly Summarized**

As previously mentioned, a property deflationist who is also a thorough-going deflationist (e.g., Horwich) needs an account of meaning to complement her deflationary account of truth. Horwich has done an impressive job of developing a use-based account of meaning that is both reductive and substantive, but that purports to be deflationary in the sense that its meaning conditions do not capture or encode truth conditions.

Although I believe that Horwich’s account appears to be plausible, ultimately, I do not think that it will work. In what follows, I will develop criticisms of it that are designed to

bring out issues that are both specific to Horwich's project and relevant to the deflationist program as a whole.

According to Horwich, the meaning of an expression is captured by that expression's basic regularity of use. Since his is not a traditional reductive account of meaning, Horwich thinks that he can have his cake and eat it to: he can combine a substantive (though not in the standard way) notion of meaning with a deflationary account of truth. As I will try to show, it is doubtful that this can be done, both as Horwich has tried to do it, and more generally.

As we have previously discussed, Horwich believes that the disquotational schema (DS)—“p” is true iff p—is conceptually basic, in the sense that its instances are accepted in the absence of argument or inference. Hence, he takes the core of deflationism to be a fact about our concept of truth. Furthermore, he takes each instance of the schema to partially axiomatize ‘true’, and he takes the totality of instances to fix its meaning. We should also note that he takes reference and the ‘true of’ (i.e., predicate reference) relation to be implicitly defined by all (non-paradoxical) instances of (R) and (PR), respectively.

(R)  $(\forall y)$  (“F” is true of y iff Fy)

(PR)  $(\forall y)$  (“N” refers to y iff  $N = y$ )<sup>167</sup> (1997, p. 100)

Finally, we note that Horwich's use theory of meaning is *substantivist*—he reduces meaning properties to non-semantic regularities of use. This should be qualified: although Horwich does believe that the meaning of a sentence should be identified with

---

<sup>167</sup> For present purposes, Horwich's motivation is unimportant. Furthermore, it, as well as his position on the other aspects of the deflationist program, is discussed in other chapters.

an underlying, non-semantic property of that sentence—a use property—he does not suppose that the properties underlying all instances of the M-schema—“s” means that  $p$ —can be explained by reference to one unified property,  $P$ , underlying all cases of meaning, as previously noted. Instead, he believes that the various instances of the M-schema, sentences like “‘s<sub>1</sub>’ means s<sub>1</sub>, “‘s<sub>2</sub>’ means s<sub>2</sub>, ..., “‘s<sub>n</sub>’ means s<sub>n</sub>” should be read as (e.g.) ‘s<sub>1</sub>’ is used in way<sub>1</sub>, ‘s<sub>2</sub>’ is used in way<sub>2</sub>, ..., ‘s<sub>n</sub>’ is used in way<sub>n</sub>.

Horwich’s use theory of meaning involves three claims. First, Horwich identifies meanings with concepts, abstract entities that compose the contents of our attitudes and the meanings of our words. Second, he holds that each word is governed by a basic regularity of use, the small core of facts that explain everything that we do with the word. Hence, he holds that meaning properties are *constituted* by use properties, and a word expresses the concept that it does in virtue of its use properties. Third, Horwich holds that for two words to have the same meaning is for them to express the same concept, which means that given two words,  $w$  and  $v$ ,  $w$  means the same thing as  $v$  if and only if  $w$  and  $v$  have the same basic regularity of use.

Although Horwich does not explain what it is for a meaning property to be constituted by a use property, given other things he says, it is not hard to divine his intention.<sup>168</sup> For Horwich, the claim that use constitutes meaning amounts to (or entails) that meaning facts can be completely explained by reference to facts about use. So, the claim that the meaning of a word is constituted by its basic regularity of use is, or, at least, entails, the claim that the basic regularity of use explains whatever meaning facts there are to be explained.

Before I turn to my central worries, I would like to consider an issue that Horwich has not addressed, one that I think make his account at least a bit hard to accept. The first worry that comes to mind regards the publicity of meaning, which Horwich endorses. It strikes me as highly implausible to assume that, for any  $w$  of mine, there is a  $v$  of yours such that the same set of basic facts explain my use of  $w$  and your use of  $v$ . Although I think that this problem holds for almost every word, it can be brought out clearly through a consideration of special sorts of words. To take an example, why think that you and I share the same basic regularity of use with respect to our moral or aesthetic terms, especially given certain social, etc. differences that exist between us? My question is not, Why think that we use our moral or aesthetic terms in the same way?, for even if we do, that is compatible with the claim that the *basic regularities* of our uses are different. Rather, my question is, Why think that the basic regularity governing my use of, say, ‘good’—the fundamental law governing the use, as Horwich says—is the same as the basic regularity governing your use? Horwich needs to say something about why we should believe that there are such shared basic regularities of use, especially in light of the skepticism that many people have regarding interpersonal sameness of conceptual role.

There is a further issue, one that arises even if we are unimpressed with the issue regarding publicity of meaning. Why believe that there are any basic regularities governing use, i.e., any “laws of use”, as Horwich (1995, 1997, 1998) calls them? It is true that we are able to adduce seeming basic regularities of use for conjunction, and the

---

<sup>168</sup> My evidence comes from Horwich’s discussion (1993) of the problems with claiming that reference is constituted by causal properties.

like, as Horwich notes, but why think that there is a basic regularity of use governing every expression type?

It is central to the deflationist line of argument that unless the inflationist can furnish a reason for thinking that the project of providing an explanatory account of truth and truth conditions can be carried through, the inflationist is not in a position to assume that we can provide an explanatory theory of truth or truth conditions at all.<sup>169</sup> However, Horwich provides us with no reason for thinking that there are such laws of use. Furthermore, it is notoriously difficult to imagine what one would look like for anything as mundane (and observational) as **Cat**. Given the wide array of linguistic contexts in which it is used, the variation in cats, etc., absent a showing of some basic regularity that explains all of our uses of ‘cat’, we have no reason for believing that these laws of use exist.

Hence, to at least some extent, Horwich is in the same boat as the substantivists: by his own rules, either he must provide us with good reason for believing that a basic regularity governing the use of each expression of the language can be uncovered, or he must abandon his position. Given the difficulty that I imagine will arise with something like ‘cat’, it is anyone’s guess whether or not such a thing exists for ‘charm’, ‘two’, ‘positron’, or any of the cases that we really care about. To this extent, then, even excepting the problems noted above, Horwich is in no better shape than are the substantialists against whom he is arguing.

Although I do think that the aforementioned worries need to be answered before Horwich’s account can be adopted, and although I believe that it will be hard, if not

impossible, for him to answer them, I will not pursue these issues any further. There are three problems and one general issue that I take to be considerably more serious, and that I will concentrate on in what follows. First, I will discuss the general issue and the two general problems regarding his use theory of meaning, and then I will turn to one more specific problem regarding the work that he believes his combined use theory of meaning/deflationary theory of truth package can do.

#### 4.2.1 The General Issue and the General Problems

Before we turn to the two general problems with Horwich's account of meaning, I would like to raise an issue. Horwich (1997, p. 100) asks us to

[c]onsider the English word "dog". Which of its non-semantic properties(if any) is responsible for its having the particular meaning it has? Or to put it another way, which property 'Ux' satisfies the reductive theory  $x$  means *dog* = Ux?

Horwich's answer is that it is some *use* property that is responsible for it having the particular meaning that it has, but we are not concerned with his positive answer quite yet. Our first question is, Why should we expect to be able to adduce a reductive theory of meaning at all?

Horwich's answer is that we need an account of meaning to explain how language can have the representational properties that it has. He (Ibid.) claims that

the main difficulty in answering this question [which property 'Ux' satisfies the reductive theory  $x$  means *dog* = Ux?] derives from our conviction that the meaning of a predicate determines what it may correctly be applied to: in particular that

$x$  means *dog*  $\rightarrow x$  is true of dogs.

---

<sup>169</sup> Arguments to this effect have been developed in Field (1986, 1995), Horwich (1990, 1995, 1997, 1998), and Leeds (1978, 1995), among others.

For it is assumed (and rightly, I think) that our commitment to this meaning-to-truth conditional places a substantial constraint on what ‘*Ux*’ might be.

The general argumentative strategy that I will employ proceeds as follows:

- (A) Per Horwich, the primary reason for being interested in getting an explanation of meaning is that we are interested in explaining “the representational powers of language”—as Horwich puts it, “How it is possible for a word—a mere sound or mark—to represent some specific aspect of the world” (1997, p. 99), which, as we have described it, is an external question; so
- (B) His goal is to show how representation is possible.
- (C) He holds that the meaning-to-truth conditionals are central to the representational powers of language; so
- (D) He must explain the putative fact that the meaning-to-truth conditionals actually hold; but
- (E) His argument, if successful—which it is not, anyway—would show only why we accept the meaning-to-truth conditionals, and not why they hold, which means that
- (F) He has not achieved his goal (in (B)); and, therefore,
- (G) He has not developed an account of meaning that satisfies the desideratum that he himself has claimed must be satisfied.

So, Horwich’s claim to be explaining how representation is possible is false, as we will see. What he actually explains, if anything, is why we accept the meaning-to-truth conditionals. If the meaning-to-truth conditionals are as central to the “representational power of language” as Horwich claims, then, *pace* Horwich, he has not explained the representational power of language.

This raises an interesting question: why should Horwich explain the representational power of language? As we have described the general semantic deflationist strategy, what the deflationist does is to divide explananda into two classes—the class of putative semantic questions that can be raised from within the language—Is this sentence true? Does this expression have the same meaning as that one?, etc.—and reflective philosophical questions that can be asked *about* our linguistic practices—Why do we ascribe meanings to expressions and sentences? What purpose is served by having a truth predicate?, etc. Once this distinction is in place, the deflationist can answer putatively external questions, without presupposing that we must provide deep or substantive accounts of the natures of the properties the substantivist assumes to be picked out by our semantic concepts.

I take it that the question, How is representation possible?, is the mother of all external semantic questions, and, for this reason, I wonder why Horwich thinks that it need be answered in the usual way. More to the point, insofar as he adduces a reductive account of meaning because he wants to answer the question about the possibility of representation, we might wonder whether or not it is really something that he needs. That is, it is unclear why the deflationist needs a reductive account of meaning, given that she really is not in the business of explaining the possibility of representation.

What the deflationist should notice, I think, is that, by taking on the task of answering questions like (58), she is buying into the substantivist's

(58) In virtue of what does this mark—**Snow**—mean *snow*?

program, which takes questions like (58) to be, in many ways, the most basic semantic questions we can ask. If they are correct about the status of such questions, then Horwich

is right to wheel out his reductive use theory, though I fear that he will be unsuccessful at showing that we can combine a more-than-deflationary account of mean with a deflationary theory of truth.

Summing up, then: it is important to note that the deflationist need not accept the substantivist's contention about questions like (58). What the deflationist should say about such questions is that their answers will be somewhat derivative: rather than reading (58) as the substantivist does, the deflationist should read it as something like (59).

(59) How does it come to pass that English speakers interpret this mark—**Snow**—as the English word 'snow', which means *snow*?

Here, we might think, an explanation in terms of use would be very handy, but the deflationist would not want a Horwichian reductive use theory of meaning; instead, she might invoke facts about our usage to explain our practice of taking, reading or interpreting certain physical objects in certain ways.

This is a brief sketch of what I think that Horwich should have said, but notice that if he had said it then he never would have offered a reductive use theory of meaning; instead, he would have offered a *deflationist* theory of meaning, one that invokes facts about use to explain our practices without, at the same time, buying into the substantivist's project of attempting to explain, in any serious sense, the possibility of representation in a natural world.

As I will argue, somewhere in the middle of his discussion of meaning, Horwich appears to have changed the subject: his stated goal was to explain the representative power of language, which, he claimed, required explaining the meaning-to-truth

conditionals; what he actually does is (attempt to) explain why we think they hold. Although it should now be clear why I think he didn't need to invoke a substantive theory of meaning, as I will show, there are good reasons for thinking that, consistent with his deflationism, he *couldn't* invoke a substantive theory of meaning. However, as will be clear, this is exactly what he needs if he is to maintain that the meaning-to-truth conditionals really do hold. Horwich cannot have his cake and eat it too.

In a footnote (1997, p. 103, fn. 9), Horwich claims that the meaning-to-truth conditionals might be explanatorily basic, which means that, though they do hold, we cannot explain why they hold. This might seem to give him the sort of wiggle room that he needs. Appearances are deceiving, however, for, given (C), above, the claim that the meaning-to-truth conditionals are explanatorily basic is tantamount to the claim that the representational power of language—how it is possible for a mere sound or mark to represent some specific aspect of the world—cannot be explained. I take it as a datum that no one—Horwich included—would accept this claim for a second; i.e., no one would both accept that language really does represent *and* accept that it is utterly inexplicable how this trick is pulled off. Hence, I take it that, if we accept the centrality of the meaning-to-truth conditions to the representational power of language and we accept, with Horwich, that the representational power of language must be explained, then we cannot take the meaning-to-truth conditionals to be explanatorily basic, unless we are willing to accept the seemingly outlandish view that the representational power of language is a basic fact about the universe.

#### **4.2.2 Our First General Problem with Horwich's Account of Meaning**

Horwich is not very clear on the goal of his project in the theory of meaning, for he accepts the following, seemingly incompatible, claims: first, that the meaning-to-truth conditionals are central to our explanation of the possibility of representation; second, that, armed with his use theory of meaning, he will explain how representation is possible; and, third, that he will not explain how it is that the meaning-to-truth conditionals hold, but only why we suppose that they do. Hence, it is unclear whether or not he will actually explain the possibility of representation at all. In fact, in his latest word on the subject (1998, 100), he notes that the claim that we could actually provide a reductive account of linguistic representation is central to an inflationist account of representation, and, therefore, not something that *he* could endorse.

Given this claim, it is utterly obscure how he can claim that “to the extent that reference and truth are ‘deflated’ it is surely going to be easier to show how representation is possible.” (1997, p. 100) Horwich is right that the thorough-going deflationist would not want actually to explain the possibility of linguistic representation. However, as I will now show, whether he wants to explain representation or not, armed with his reductive account of meaning, he has all of the machinery necessary to provide a reductive account of representation, and, therefore, of truth. Since I believe that this result does not turn anything unique to Horwich’s account, but is, instead a consequence of the combination of a substantive theory of meaning with acceptance of the meaning-to-truth conditionals, this raises serious questions about whether or not the deflationist about truth can adopt a theory of meaning at all. Furthermore, I believe, it makes more pressing our question, Why would a deflationist about truth develop a reductive account of meaning?, which is a question of justification, rather than of motivation.

As a deflationist about truth, Horwich denies that we can (or must) come up with a principle of the form (IT), which characterizes inflationist accounts. He also denies

(IT)  $x$  is true  $\leftrightarrow x$  is P

that we can (or must) adduce similar accounts for reference, e.g., (IR)

(IR)  $x$  refers to  $y \leftrightarrow xRy$

or the “true of” relation, e.g., (ITO)

(ITO)  $x$  is true of  $y \leftrightarrow xCy$ .

However, as an inflationist (of any sort) about meaning, he does hold that there is a substantial, non-semantic use property,  $U$ , that constitutes the meaning of a predicate  $F$ .

Hence, he accepts that, for each predicate of the language,  $F$ , there is a non-semantic meaning property  $U$  such that (IM) holds.

(IM)  $x$  means  $F \leftrightarrow Ux$

Consider the fact that, per Horwich’s reductive account of meaning, an instance of (IM) holds for each meaningful entity,  $x$ . Given that we can explain—in reductive terms—what it is for  $x$  to mean  $F$ , we can explain, in reductive terms, what it is for  $x$  to be true of  $y$ : for  $x$  to be true of  $y$  is for there to be some property  $F$  such that  $x$  means  $F$  and  $y$  is  $F$ .

That is, given (IM), it is inevitable that we will be able to define the ‘true of’ relation by (HM). Hence, though, in principle, Horwich denies that we must give reductive accounts

(HM)  $xCy \leftrightarrow (\exists F) x$  means  $F$  &  $y$  is  $F$

of our semantic relations—i.e., as in (IT), (IR) and (ITO)—he has provided us with the materials necessary to get a reductive account of those semantic relations. Given that we have it that ‘ $x$  means  $F$ ’ can be reduced to  $Ux$ , we are guaranteed that the right-hand side will be free of semantic machinery, from which it follows that Horwich can give a

reductive account of representation, after all. What this means is that Horwich's substantive account of meaning has bought him a substantive account of the "true of" relation, and, of course, the point generalizes to reference and truth, since, as Horwich shows (1998, p. 109), all three can be inter-defined. What this means is that Horwich's inflationist view of meaning has resulted in an inflationist account of representation, and, hence, of truth conditions.

How might Horwich respond to this charge? First, Horwich would balk at the claim that he has bought himself a substantivist (inflationist) account of representation and truth; after all, as Horwich himself points out, his account of meaning is, in some sense, deflationist: he does not reduce meaning to some single, non-semantic relation. Given the form of (HM), the same would be true of representation. Hence, even if he is committed to something like (HM), he might still argue that he has provided a deflationist account of representation.

Horwich's claim to have produced a (somewhat) deflationary account of meaning is suspect, given the way in which we are understanding contemporary deflationism. Of course it is more liberal than the orthodox accounts of meaning (i.e., those that try to find some one non-semantic categorial base to which meaning properties can be reduced), but it does not follow that it is not substantivist or inflationist,<sup>170</sup> unless we also suppose that only those theories that reduce all meaning properties to some one non-semantic property are substantivist. I can see no reason for doing that. In fact, I think that it is obvious that the account of representation that Horwich ends up with is substantivist (or inflationist)

---

<sup>170</sup> In this section, instead of just using 'substantivist', I will be using 'substantivist'; and 'inflationist', in order to distinguish the claim that Horwich would not accept—that, as things stand, he is committed to an

by Horwichian standards. Since our way of understanding contemporary deflationism comes to us largely through Horwich's own work, it follows that (HM) and the account of representation that, per Horwich, it grounds, if true, would be inflationist by Horwichian standards.

If we are right in our claim that Horwich's account of meaning buys him something like (HM), then, whether or not we characterize his account as inflationist, we can see that it is at least incompatible with the deflationary account of truth that it was developed to complement. This puts Horwich in an awkward position, as we have noted, but this is not the most significant fact about our argument. Notice that our derivation of (HM) from Horwich's reductive account of meaning did not depend on the specifics of his account. If we are right that we can get (HM) from (IM), then two things follow. The first thing that follows is that a deflationist about truth cannot explain the representational properties of language. Put another way, what follows is that a thorough-going deflationist about truth must be a deflationist about meaning. There are reasons for worrying about this claim, for at least some of the reasons for caring about meaning—I think, for example, of claims regarding communication, although there are myriad examples from which to choose—have not regarded truth at all, or, at least not directly. Even if this claim can be undercut, as I will note later, her commitment to deflationary meaning places on her a rather substantial burden: she must provide an account of the role of 'means that' that is akin to the account she provides for 'true'. That is, if she is going to deny the existence of objective meaning properties, i.e., if she is going to deny to the meaning predicate its usual function in the language—viz., to pick

---

inflationist account of representation—from the claim that he would accept—that he has developed a non-

out a property—then she must say what it does, and why we have it. This is a point that seems to have been missed by the deflationist, as we will see.

The second thing that follows, which is related to the first, is that if, as Soames has argued, we need a substantive account of meaning, then thorough-going deflationism is false.<sup>171</sup> Since, by our argument above, the proponent of thorough-going deflationism<sup>172</sup> must deny that we can provide a substantive theory of meaning (or representation), what this means is that the explanatory burden that is placed on the deflationist is somewhat greater than she has thought. It is not enough to point to the failures to adduce adequate accounts of meaning or representation, as, e.g., Horwich (1995, 1997, 1998) and others have done, what must be shown is either (i) that there is a principled reason—N.B., such as the indeterminacy of translation, if only that could be made to work—for believing that it is *impossible* to develop a substantive account of meaning or representation; or (ii) that we have no need for a substantive account of representation, i.e., that everything that we need meaning for can be had without the supposition that there is a substantive property of meaning. To my knowledge, this issue has not yet been adequately addressed by the deflationists.<sup>173</sup>

---

inflationist account of meaning.

<sup>171</sup> Incidentally, this is why I express skepticism about Soames' claim (1995, 1997) to be developing a deflationary account of truth. Soames wants to combine an inflationist—in fact, a direct reference—account of meaning with a deflationist account of truth. Soames' deflationist account is suspect for a variety of reasons, the most important of which is that he believes that truth has an important explanatory role to play in his version of the meaning-to-truth conditionals that we are currently discussing. I should note that, in claiming that his view is suspect as a deflationary account, I am not claiming that it is wrong—I am merely claiming that it might be wrongly characterized.

<sup>172</sup> Just to remind the reader, a thoroughgoing deflationist view of truth is one that is deflationist about truth for all putative truth bearers. Hence, a view like Soames' (1984, 1997, 1998) is deflationist—because it is deflationist about propositional truth—but, given our argument, it is not a thoroughgoing deflationist view, since it is not deflationist about what it is for a sentence to express a proposition, i.e., about meaning.

<sup>173</sup> I will discuss this issue in more detail later, but for now I should point out that in his excellent paper, "Deflationary Views of Meaning and Content," which, by the way, is primarily concerned with the deflationary theory of *truth*, anyway, Field does not say what we need meaning for. What he does say, and

### 4.2.3 Our Second General Problem: Horwich and the Problem of Intentionality

What we have established thus far, then, is that Horwich has not satisfied (B), which means that he has not shown how representation is possible. As we will show now, he has not satisfied (D), either. If we are correct, then it follows that Horwich has not accomplished what he has set out to do. In fact, if we are correct then it appears to follow that what he has set out to do simply cannot be done.

So, the second problem that I want to consider regards Horwich's attempt to derive the meaning-to-truth conditionals from his use theory of meaning, together with a deflationist account of truth. As mentioned previously, Horwich's claim to have solved the problem of intentionality was somewhat exaggerated—at one level, he attempts to “solve” it by arguing that it is not something that the deflationist has to worry about. In that sense, he attempts to *dissolve* it. In another sense, however, he may have solved some aspect of it: if his derivation goes through, he has explained why we accept the meaning-to-truth conditionals, and, for a deflationist, doing that is the surrogate for the allegedly “deep” solution sought by the inflationist. However, as I will show, it is doubtful that he has even done that much: the problem of intentionality, in both its inflationist and its deflationist guise, remains.

The problem of representation or aboutness is the problem of saying how it is that a sentence or word could mean or be about something. Hence, a solution to the problem would provide us with the tools necessary to say what it is for a sentence to be true, for a predicate to be true of something (or some things), and for a term to refer to what it does.

---

what is very informative, is that the deflationist can mimic substantivist accounts of meaning, by making deflationary, non-truth conditional meaning and content both externalist and social. Of course the

Horwich cites, as the heart of the problem of representation, the fact that the meaning of a sentence (or a predicate, etc.) determines its extension, i.e., a truth value (or a set, etc.). Per Horwich, we accept the general fact about the relation between meaning and truth, and we accept the particular facts that are its instances—the meaning-to-truth conditionals—and these are the central data that need explaining.<sup>174</sup>

As we noted previously, if it is a fact that the meaning of a sentence or a predicate determines<sup>175</sup> its extension, then this must be explained; by contrast, if it is not a fact that meaning determines (or provides a necessary condition for) extension, but if it only appears to be one—if, for example, we draw an illegitimate conceptual connection between meaning and truth—then it needs to be explained away.<sup>176</sup> An adequate account of meaning and its relation to truth must either explain how linguistic representation is possible, or it must explain why we think that it is possible.

According to Horwich (1997, p. 101), then, if the deflationist is to fulfill her obligation to explain everything that there is to explain about the conceptual and explanatory roles of truth, she must explain (or, derive) the meaning-to-truth conditionals, and, since she is a *deflationist*, she must do so without invoking any heavy-duty semantic machinery. Here is Horwich (1997, p. 102), on how the deflationist should derive the meaning-to-truth conditionals:

---

inflationist have assigned to meaning an explanatory role—the role of explaining truth and truth conditions. Needless to say, the deflationist could hardly agree with this claim.

<sup>174</sup> Following Horwich, I will concentrate on predicate reference, the “true of” relation.

<sup>175</sup> We could weaken this claim with no harm done to my central point if, say, we held that extension was partially determined by context, as Katz (1997) does.

<sup>176</sup> Of course, this only holds for one who is not providing a revisionist account of truth, meaning, reference, and the like. For one providing a revisionist account, if, as a matter of fact, there is no explanatory relation between meaning and truth, then there is nothing to be explained.

For insofar as the disquotational schema is conceptually basic, our natural route to the meaning-to-truth conditional

$x$  means **dog**  $\rightarrow$   $x$  is true of dogs

is to derive it from the three principles using the following reasoning.

(60) “dog” means **dog**.<sup>177</sup>

(61) “G” means the same as “dog”  $\rightarrow$  (y) (Gy  $\leftrightarrow$  y is a dog)

(62) (y) (“G” is true of y iff Gy)

From (61) and (62) we get (63)

(63) “G” means the same as dog  $\rightarrow$  (y) (“G” is true of y iff y is a dog)

And this, given (60), implies (64)

(64) “G” means **dog**  $\rightarrow$  (y) (“G” is true of y  $\leftrightarrow$  y is a dog)

from which we find it natural to infer—generalizing from the predicates “G” that we understand, to all predicates—(65)

(65) (x) ((x means **dog**)  $\rightarrow$  (y) (x is true of y  $\leftrightarrow$  y is a dog)).

In this way, Horwich thinks we can justify our acceptance of the meaning-to-truth conditionals. Since he also believes that, again, from a deflationary perspective, the meaning-to-truth conditionals capture what there is to capture regarding the representational character of language, he believes that, in so doing, he has provided the means that are necessary to solve the problem of intentionality. This claim turns on the further claim that, from the above derivation, we can derive general conclusions, for that is what a solution to the so-called “problem of intentionality” demands. Furthermore,

---

<sup>177</sup> In this dissertation, as previously mentioned, I use boldface for concepts and italics for properties. Horwich uses italics for concepts, but, for the sake of consistency of presentation, I will not follow him.

and at least *prima facie*, his derivation relies on wholly acceptable premises: (60) follows from his use theory, whereas (61) is taken to be explanatorily fundamental and (62) is understood as *conceptually basic*.<sup>178</sup>

There are numerous criticisms that could be made of this derivation. I will mention what I take to be the most important two: first, contrary to Horwich's claim, we cannot derive general conclusions from this derivation; and second, the premises of the derivation are not acceptable.

We will take the first criticism first. To be sure, Horwich tells us that we find it natural to move from (66) to (67), but that does not amount to the *derivation* of (67);

(66) "G" means **dog**  $\rightarrow$  (y) ("G" is true of y  $\leftrightarrow$  y is a dog)  
 (67) (x) ((x means **dog**)  $\rightarrow$  (y) (x is true of y  $\leftrightarrow$  y is a dog))

Horwich has promised more than his argument can deliver. But take that point, for, Horwich has not even established that any predicate that means **dog** is true of (or applies to) dogs. What his argument yields—(66)—is a conclusion about a specific predicate, "G", to the effect that if it means **dog** then it is true of dogs.

Maybe I am being too hard on Horwich; after all, we can transform (66) into a generalization quite easily. In order to do this, let's replace (66) with (68).

(68) For all English predicates, "G", "G" means **dog**  $\rightarrow$  (y) ("G" is true of y  $\leftrightarrow$  y is a dog)

Of course, (68) does not suffer from the problem of (66), but it, too, is problematic, at least for Horwich. The reason is that (68) fails to account for the general fact that

---

<sup>178</sup> Actually, Horwich vacillates on the status of (61), sometimes calling it conceptually basic, sometimes calling it explanatorily fundamental.

Horwich is trying to account for, namely, that any predicate, of any language, that means the same thing as the English expression ‘dog’ is true of the dogs.

Again, one might think that I am not being fair to Horwich, for it appears that the problem just noted can be solved if we do not restrict our quantifier as we did, i.e., if we opt for (69) over (68). However, it is easy to see that this emendation is a non-starter.

(69) For all predicates, “G”, “G” means **dog**  $\rightarrow$  (y) (“G” is true of y  $\leftrightarrow$  y is a dog)

The problem with opting for something like (69), as I see it, is that Horwich’s argument depends on the disquotational schema for predicate reference, i.e., premise (62), and that is not available for predicates of other languages.<sup>179</sup> So, for example, the Hebrew word ‘sus’ means the same thing as ‘horse’.<sup>180</sup> However, (70) makes no sense. Of course, (71)

(70) (y) (“sus” is true of y iff y is a sus)

(71<sup>E</sup>) (y) (“sus” is true of y iff y is a horse)

makes sense, and, in fact, is true, but it is not an instance of the predicate reference (‘true of’) disquotational schema; it is not, as Horwich says, “conceptually basic.”

What this means, then, is that the conclusion to which Horwich claims we can generalize—i.e., the limited conclusion about all predicates that mean the same thing as “dog”—cannot be derived from his premises. What can be derived from his premises?

Well, it appears that (69), which is weird, anyway,

(69) For all English predicates, “G”, “G” means **dog**  $\rightarrow$  (y) (“G” is true of y  $\leftrightarrow$  y is a dog)

---

<sup>179</sup> It is crucial that, in (62), “G” ranges over predicates, rather than concepts. Since Horwich is interested in explaining the representational properties of language, of course, then “G”, in (62), does range over predicates, rather than concepts.

<sup>180</sup> This would be considerably more dramatic if my computer enabled me to type out the Hebrew. For the purposes of this case, pretend that the English transliteration is Hebrew.

insofar as it requires quantification over name and predicate places, can be derived from his premises, but that hardly shows that Horwich has captured everything that needs capturing. Although Horwich may be able to provide derivations like (69) for every predicate of the language, this does not add up to the generalization that Horwich seeks.

Before we turn to the second problem, we should pause for a moment. Horwich's goal is to explain why we accept the generalized meaning-to-truth conditional, at least that is the goal as stated. In his attempted derivation of this conditional, Horwich adverts to a putatively obvious fact: that it is natural to generalize the result of the derivation for 'dog', and for other expressions that we understand, to all predicates, even to those of foreign languages. I have come to believe that this claim is illegitimate. There are two reasons for this. First, advertng to the naturalness of the generalization is not explanatory. If the goal was to explain why we accept the generalized meaning-to-truth conditional (or all instances of it, including those for foreign expressions and expressions that we do not understand), it is of no help to be told that we accept it because it is natural for us to do so. Indeed, that we accept it is not at issue.

Second, what Horwich needs to explain is *why* we accept the generalized meaning-to-truth conditional (or, at the very least, why we accept all instances of it). The derivation for "dog" is intended to be a rational reconstruction of the sort of reasoning that would lead us to accept the meaning-to-truth conditional for that expression. If we would accept the meaning-to-truth conditionals for predicates of other languages and for predicates that we do not understand (still understanding this as a rational reconstruction of the sort of reasoning that we would go through), we must do so by first generalizing from the instances that we do accept, and then applying the generalization—being

disposed to apply the generalization, or what have you, to the foreign and unfamiliar cases. The “naturalness” comes in at this second stage, if anywhere: we find it natural to apply *the generalization* to foreign or unfamiliar cases. If this is correct, then, if Horwich’s account is to be adequate, he must show how we can derive the generalization of the meaning-to-truth conditionals, and, as we have previously discussed, this is something that he cannot do.

Let’s turn to the second problem that I said that I would mention. The second problem regards Horwich’s second premise, which does quite a bit of

(61) “G” means the same as “dog”  $\rightarrow$  (y) (Gy  $\leftrightarrow$  y is a dog)

work in the argument. In fact, (61) provides, or, at any rate, appears to provide, the words-world connection that we expect an explanation of the meaning-to-truth conditionals to forge, and without it, or without something of equivalent strength, Horwich would not be able to derive even his limited conclusion about “dog”. However, we are in our rights to wonder whether or not (61) is an acceptable premise.

The problem is that (61) seems to be of the same form as the meaning-to-truth conditionals, which suggests that if the latter needs explaining, so does the former. If (61) is in no need of explanation then what is so special about the meaning-to-truth conditionals such that they (or our acceptance of them) are in need of explanation? Further, if they are in need of explanation, why isn’t (61)? Horwich never tells us.

This points to a larger problem, I think. Horwich claims that, of the premises in his argument, (60) is supplied by his use theory of meaning; (61) is explanatorily fundamental, meaning that we cannot (and need not) explain why it holds; and (62) is conceptually basic, meaning that we are inclined to accept it without inference or

argument.<sup>181</sup> Horwich holds that the meaning-to-truth conditionals might be *explanatorily* basic, but that they are not conceptually fundamental (or basic), as they are derived from (60), (61) and (62). The problem, as I see it, is with the claim that (61) is explanatorily fundamental.

Horwich (1997, pp. 102-103) tells us quite a bit about what follows from the fact that we accept (61), but he does not tell us where he thinks that our acceptance of (61) comes from, and this is a problem. To be sure, he claims that (61) is explanatorily basic, but, as previously mentioned, this means simply that we cannot (and do not need to) explain why it holds. In particular, this does not absolve him of the responsibility of explaining why we accept it. Of course, he might have slipped—he might have meant that it was conceptually basic—but I believe that even if that is what he meant, it is an open question whether or not he is correct.

The problem is not that his explanation of the derivation is incomplete because he has not explained why we accept one of its key premises. The problem, as I see it, is that the best explanation for our acceptance of (61) will presuppose our acceptance of the relevant instance of the meaning-to-truth conditional, or at least, of something as strong. In order to see this, we should first note that, given our acceptance of (62), (61) is equivalent to (72).

(72) “G” means the same as “dog”  $\rightarrow$  (y)(“G” is true of y  $\leftrightarrow$  y is a dog)

---

<sup>181</sup> Horwich actually makes different pronouncements about (62). In his 1997 paper, he claims that it is explanatorily basic; in his 1995 paper, he holds that (something like) it is conceptually basic. I am taking his later claim to be the one that he currently accepts. As should be obvious, however, nothing in my argument turns on that. If he takes it to be conceptually basic, well, as we will show, there is good reason for thinking that he is mistaken.

Because my point is more easily seen through a consideration of (72), we will work with (72). If we consider (72), and we consider the fact that we accept (73), then we will (73) (y)(“dog” is true of y  $\leftrightarrow$  y is a dog) recognize that our acceptance of (61) presupposes that we accept that sameness of meaning preserves sameness of truth conditions. However, if we accept that sameness of meaning preserves sameness of truth conditions, we must already accept that there is a connection between meaning and truth conditions, a connection like the one established in the meaning-to-truth conditionals. Hence, I cannot see that (61) is an acceptable premise in Horwich’s argument.

As mentioned previously, the most charitable way to understand Horwich’s argument requires accepting that he has radically overstated his goal: he is not really trying to explain how representation is possible; rather, he is trying to explain why we accept the meaning-to-truth conditionals. This means that, at best, what he could explain is why we think that representation is possible, not why, in fact, representation is possible. However, if we are right, he has not succeeded at showing even that.

This raises an important issue, I believe. Horwich starts out telling us that “to the extent that reference and truth are ‘deflated’ it is surely going to be easier to show how representation is possible.” As we have shown, Horwich’s actual goal, which he did not reach, was more modest than his implied goal, for his actual goal was, in effect, to explain in a suitably deflationist way why we accept that representation is possible. Since Horwich has not done that, the question arises as to whether or not he thinks that it is something that should be done—ought the deflationist to explain how representation is possible?

Suppose, as Horwich seems to think, that the deflationist really cannot explain why the meaning-to-truth conditionals hold. This raises two possibilities: either the meaning-to-truth conditionals are explanatorily basic, meaning that they hold, but that we cannot point to other facts about meaning, truth, or anything else to explain why they hold; or, though we accept that they hold, in fact, they do not. When the first claim is combined with Horwich's other claim, to the effect that the meaning-to-truth conditionals lie at the heart of the problem of intentionality, what results is highly problematic, and not something that I think he would take seriously, for it is the claim that representation is possible, though utterly inexplicable, which places him in a highly inflationist camp, to say the least. When the second claim is combined with Horwich's claim about the role of the meaning-to-truth conditionals in representation, the result is a non-factualism about representation. This is a position that deflationists might be happy with, but it is hard to see how the way to show "how representation is possible" is by adopting a non-factualist attitude towards it.

In conclusion, I should note that one of the reasons that a careful consideration of Horwich's account of meaning is important is that the deflationist, by severing the explanatory connection between meaning and truth, creates for himself a substantial burden: he must explain, in non truth-theoretic terms, what makes it the case that a sentence or expression has the meaning that it has, or he must explain why we should be meaning skeptics. There is overwhelming evidence that Horwich is not sympathetic to meaning skepticism, which means that he has to develop an adequate, non truth-theoretic account of meaning. The upshot of this chapter is that he has a considerable amount of work to do if he is to realize that goal.

### 4.3 Deflationist Meaning

What we learned in the last section, among other things, was that the deflationist had two options with respect to Horwich's *problem of intentionality*, either she accepts the meaning-to-truth conditionals as explanatorily basic, which option seemed absurd to us, or she go non-factualist and deny that, strictly speaking, there was any problem of intentionality. If she takes the latter option then, like Horwich, she might try to explain why we accept the meaning-to-truth conditionals, but it is not at all clear that she must, for she can take them to be conceptually basic. If she does then she might argue, in a way that I imagine Field would, that it is just a fact about us and about our conceptual schema that we accept that, to take an example, if my 'dog' means *dog* then it is true of the dogs.<sup>182</sup> On this view, we might explain why we accept the meaning-to-truth conditionals by pointing to the fact that they are trivial consequences of the meaning of 'means that' and 'true', and that, in this sense, they are *conceptually basic*. We might also hold that, since we deny that there are meaning facts,<sup>183</sup> they are not explanatorily fundamental since, strictly speaking, they are not factual (or, true; or, whatever). This, then, is what the deflationist should say about the meaning-to-truth conditionals: they are conceptually basic, though they are neither explanatorily fundamental, nor explanatorily non-fundamental (i.e., in the sense that they admit of explanation).

#### 4.3.1 A Deflationist Theory of Meaning?

As we saw from our discussion of Horwich's account of meaning, the deflationist about truth cannot accept a substantive account of meaning out of which, together with

---

<sup>182</sup> I would like to thank Jared Blank for convincing me that this is most likely the option that Field would take, and forcing me to consider it more seriously.

the meaning-to-truth conditionals, a substantive account of truth can be reconstructed. What she must adopt, then, is a non-substantive, or deflationist, account of meaning, but what is a deflationist account of meaning?

There are numerous suggestions in the literature about what a deflationist account of meaning might amount to, and it is not my concern to survey all of those. The sort of account of meaning that our deflationist requires is one that is deflationist about the role of truth conditions in meaning: whatever meaning is, it cannot as one of its constituents truth conditions *under any description*.<sup>184</sup> I think that the best way to think about a deflationary theory of meaning is as a two-factor theory minus the truth-theoretic factor, which means that, on a deflationary conception of meaning, meaning is, or is at least, a matter of conceptual role, though it can include other things—Field mentions indication relations—as well.

We know why truth conditions cannot play an important role in meaning, but it is not yet clear what that means. What it means is that meaning is a not factual affair, which is to say that we cannot talk about *the* meaning of an utterance, or the fact that a sentence has the meaning that it has. Put in a slightly different way: to claim that truth conditions do not play a role in our theory of meaning is to say that there is no objective fact of the matter as to what an expression, phrase or sentence means. On the deflationist account, we can only talk about the meaning that an utterance has for me, which amounts (more or less) to the conceptual role that the utterance has for me, i.e., how I understand

---

<sup>183</sup> As well as truth facts, though that point does not need to be made here.

<sup>184</sup> Cf. Field, 1994.

it, where to understand an utterance does not involve correlating it with a proposition.<sup>185</sup>  
 This comes out in Field's discussion of translation. According to Field (1994b, pp. 406-407, n.2)

... it doesn't make sense for X to ask whether one of Y's sentences is true *as Y really intended it*, but only whether it is true *as X understood it*. This isn't to deny that translations can be *better or worse for various purposes* (and that empirical discoveries about the production of an utterance are relevant to deciding how good a given translation is); but it is to deny that talk of *correctness* of translation makes good sense. I take radical deflationism [N.B., of the sort that Field favors] to involve not just the claim that 'synonymous' (applied interpersonally) is vague; the claim rather is that it should be understood as not a straightforwardly factual predicate at all, but as an evaluative predicate (less misleadingly rendered as 'is a good translation of.'). (All italics original)

The reason that synonymy should be understood as not a straightforwardly factual predicate is because meaning is a not straightforwardly factual matter; what makes sense is not what an utterance *means* (full stop), but, rather, what it means-for-me, what conceptual role it has for me, or, what amounts to the same thing, how I understand it.

If this is all that we can mean by 'mean' (so to speak), then of course our notion of meaning will not be truth conditional. However, I think that we need to be careful.

Whereas Horwich was inclined towards a non-truth conditional theory of meaning, he got waylaid because he attempted to provide a *reductive* account of meaning. I think that the

---

<sup>185</sup> This notion of conceptual role is highly programmatic, and I intend to leave it that way, as it is not my intention to develop a deflationary theory of meaning. In particular, although Field wants to show that, in

deflationist would be ill-advised to attempt to claim that we can offer a substantive (i.e., reductive) analysis of meaning in terms of assertibility conditions, conceptual roles, or what have you, i.e., one according to which we can derive biconditionals of the form  $\sigma$  means that  $p$  iff ..., where ‘...’ is filled in by a condition (or set of conditions) both necessary and sufficient for meaning. Instead, I think that we should take the deflationist about meaning to be giving us an informal gloss (or, as Field might say, a *rough heuristic*): for a person to say what an expression, phrase or utterance means is for her to say how she understands that expression, phrase or utterance.

To the extent that the deflationist about meaning denies that there is any objective sense in which an expression, phrase, or utterance has a meaning, she is non-factualist about meanings. Since she must deny that there is any objective sense in which an expression, phrase or sentence has a meaning, one of the cornerstones of deflationism about meaning (N.B., in the sense that we intend it) is, therefore, non-factualism about meaning.

#### 4.3.2 A Lacuna in the View

There is a gaping hole in the deflationist account of meaning, which account I have just sketched, more or less. My claim is not that the lacuna cannot be filled, but only that, until it is, there is both a motivational and a justificatory problem with the view. I will address each in turn. First, the motivational problem. We have made clear why the advocate of a deflationist account of truth would want—in fact, *need*—to adopt

---

fact, the deflationism can capture the sense in which meaning is both externalist and social, I will not go through that here.

a deflationary theory of meaning, but, as of yet, we have not too much motivation for the claim that the deflationist about truth should adopt a deflationary view of meaning, unless, of course, one accepted Quine's arguments.<sup>186</sup> Given the non-factualist stance, one might think that an argument is needed for why we should accept that, contrary to what we would have thought, and contrary to how our linguistic practice would have us believe, we should accept a deflationary theory of meaning. After all, as things stand, we can talk about *the* meaning of an utterance; about two people meaning the same thing, in uttering what they do; about the translation from Spanish into English of 'perro' as 'cat' as being not just worse than the translation of 'perro' as 'dog', but as being actually incorrect; etc.

I believe that the motivation problem is intimately connected with the justification problem, and that it is really with respect to the latter that there is lacuna. In order to bring out that problem, we should return to the deflationary theory of truth, and, in particular, to the feature of it that allows the deflationist to deny that we need provide a reductive account of the property of truth. According to the deflationist about truth, attempts to provide substantive, reductive analyses of truth are misguided. The reason they are misguided is that, as the deflationist sees it, the substantivist makes the unwarranted inference from the claim that there is a property of truth to the claim that there is a complex, naturalistic, or (at least) non-disjunctive property of truth, i.e., that truth is the sort of property for which we should expect a reductive analysis.<sup>187</sup> In this way, the deflationist would argue, the substantivist arrives at the faulty conclusion that

---

<sup>186</sup> One might argue that Schiffer's (1987) arguments establish that the deflationist should adopt a deflationary view of meaning. I leave this as an open possibility.

<sup>187</sup> Horwich (1990, pp. 38-39) is explicit and admirably clear on this point.

truth is a *thick* or *substantive* property, and that it is the philosopher's job to uncover its non-semantic, underlying nature.

The claim that truth is a thick or substantive property is not very clear, but we can clarify it in a way that conforms to the writing of various deflationists as follows: a thick or substantial property is one that has a causal-explanatory role in either a mature scientific theory, a facet of our folk practices, or both. If a property is thick or substantial, then, in order to explain our use of it, or for some other causal-explanatory purpose, we must give a reductive analysis of it. What if a property is not thick or explanatory? In that case, the deflationist denies that we must give a reductive analysis of it. However, it does not follow from this that we needn't give any analysis of it. If a property is thin or insubstantial, then the predicate that expresses it does not play a role in our language the explanation of which requires us to provide a *reductive* analysis. Instead, an alternative sort of account is required.

As the deflationist sees it, the property of truth is not a thick, substantial property, from which it follows that it is not a property whose nature is of any philosophical concern, i.e., not a property for which we should expect a reductive analysis to be forthcoming. The reason is threefold: first, our ordinary concept of truth (per Horwich) is *as* of a thin, insubstantial property, which follows from the fact that possession of the concept is exhausted by one's disposition to accept, without inference or argument, instances of the disquotational schema. Second, the property of truth plays no causal-explanatory role in any of our theories; and third, the explanatory role of the property of truth is exhausted by the logical role of the truth predicate, i.e., its role as a device of generalization. This has led (e.g.) Hartry Field to claim that, if we can think of the truth

predicate as expressing a property at all, we can think of it as expressing a *logical* property.<sup>188</sup> According to the deflationist, then, we have said everything there is to say about the property of truth, when we have uncovered the linguistic role of the truth predicate.

The point of this digression has been to dispel the illusion that, as the deflationist should see it, if a particular property does not admit of a reductive analysis, it does not admit of any analysis whatsoever, for the claim that certain of the predicates of our language express properties that do not admit of reductive analysis is part of a larger view, one according to which what kind of property a predicate expresses depends in important ways on the role or function that the predicate plays in the language.<sup>189</sup> Hence, the claim that a predicate, *P*, expresses a thin or insubstantial property, *P*, entails that the “nature” of *P* can be exhaustively accounted for by the provision of an account of the role that the predicate, ‘*P*’, plays in our language, or in one or another of our explanatory practices.

So as to bring this out further, I will consider one other famous attempt to show that a familiar and seemingly thick property is actually thin and insubstantial.<sup>190</sup> In his very important paper “Reference Explained Away,” Robert Brandom (1984) argues that ‘refers’ does not express a thick or substantial property, and that, therefore, causal theorists and their opponents, all of whom are trying to reveal the nature of reference, are on the wrong track. According to Brandom, *reference* is a thin or insubstantial property,

---

<sup>188</sup> Horwich (1990) reports Field as having said this; see Horwich, 1990, p. 39.

<sup>189</sup> Of course some deflationists, notably Field, would balk at my talk of properties. I hope it is clear, to the extent that nominalist recastings are possible, that the claim that I am making on behalf of the deflationist admits of a nominalistic reconstrual.

<sup>190</sup> I discuss this case for heuristic reasons, and, in particular, not because I accept the account proffered.

which means that it is not the sort of property for which we should expect a reductive analysis. Instead, we can say everything that needs to be said about the “nature” of reference by explaining its linguistic function, which, Brandom tells us, is to construct complex anaphoric pro-forms.

Brandom claims that the primary function of ‘refers’ is revealed in cases like the following.

(74) Fred says to Lucy: ‘Ethel is upset at you.’

(75) Lucy says later: “The person to whom Fred referred is upset at me,”

where ‘the person to whom Fred referred’ is anaphorically dependent on ‘Ethel’ in the way that ‘she’ is anaphorically dependent on ‘Ethel’ in (76).

(76) Ethel came home and she found Lucy in bed with Fred.

Brandom is offering us an account of the linguistic function of the reference predicate.

He notes that

The anaphoric approach will not tell us how to understand sentences such as  
Reference is a physical, causal relation.

The reason is clear. On the anaphoric account although ‘...refers to...’ plays a syntactically relational role, its semantic role is anaphoric and pronominal, rather than relational. Philosophers have misconstrued the plain man’s use of ‘refers’... Such a mistake is of a piece with the search for objects corresponding to each expression that syntactically plays the role of a term. (1984, pp. 487-488).

Brandom claims to have identified an important difference between the role of, what he calls, ‘reference talk’ and other, ‘descriptive’ discourse, which leads him to deny that we have any need to give, or reason to expect, a reductive account of reference.

The purpose for our discussion of Brandom should be clear. Brandom's work brings out the sense in which part of what it is to adopt a deflationary attitude towards a notion (or family of notions) involves acceptance of the claim that we can explain why it is that we have such a notion (or family thereof)—the concepts of truth or reference, to take two examples—without, at the same time, assuming it to have an underlying nature the revelation of which is crucial if we are to understand the notion and if we are to account for the role it plays in one or another of our explanatory projects. This distinguishes deflationism about X from substantivism about X, but it also distinguishes deflationism about X from eliminativism or skepticism about X.<sup>191</sup> The Quinian view of semantics—at least one of the official Quinian lines—is skeptical about the prospects of rendering respectable the concepts employed in the theory of meaning, and, because of this, it is eliminative with respect to at least some aspects of meaning theory.<sup>192</sup> Quine does *not* (or, not *just*) offer a deflationary view of meaning, or meaning-theoretic semantics, showing it to be less substantive, ontologically committal, or explanatory than philosophers—though probably not the folk—would have thought. Rather, Quine offers arguments<sup>193</sup> intended to show intensionalism to be wholly confused, and, on the basis of those arguments, he rejects intensional concepts and intensional claims altogether.

The deflationist does not express skepticism about the concepts that populate the domain that she aims to deflate, and she does not want to eliminate them or show them to be involved in nothing more than “second-rate” explanations. Rather, she wants to show

---

<sup>191</sup> I believe that Horwich's failure clearly to distinguish deflationism from eliminativism is responsible for his adoption of a substantive use-theory of meaning.

<sup>192</sup> This could be a bit of an overstatement, since, strictly speaking, the Quinian replaces robust meaning with the more etiolated (and indeterminate) notion of stimulus meaning.

<sup>193</sup> For example, in “Two Dogmas of Empiricism”.

that there is really much less to the constellation of concepts than we would have thought, and that the need to provide deep or substantive reductive analyses of them simply does not arise. This means that if the deflationist about truth really adopts a deflationary view of meaning, then she must show that *meaning* is not a thick or substantive property, which means that she must show that, contrary to what at least some philosophers have thought, the “nature” of meaning can be captured by the function playing in our language or our practices by the meaning predicate.<sup>194</sup>

Deflationists like Field, Leeds, and Horwich have not tried to show, as they did in the case of truth, that the explanatory role of the property of meaning is exhausted by the function of our concept of meaning. In fact, they have said almost nothing at all about the function of **meaning** (and meaning ascriptions). Hence, they have not yet proffered anything like a complete *deflationary account of meaning*. In order to provide a complete deflationary account of meaning, they would have to provide an account of the linguistic role of ‘means that’, and they would have to show that, armed with this, we can say everything that need be said about the nature of *meaning*.<sup>195</sup> Indeed, Field (1994a, p. 254) purports to be deflationist about “the locutions “means *that*” and “has the content *that*”. However, on at least one way of understanding this claim, to be deflationist about

---

<sup>194</sup> Lest the reader suppose that I am ascribing too narrow a view to the deflationist, I should note that explaining the nature of meaning by reference to the linguistic role of the concept of meaning might not be as straightforward as it is in the case of truth. For example, it might involve the normative implications of meaning ascriptions, or some other “non-grammatical” function.

<sup>195</sup> I should note, incidentally, that I consider it an open question whether or not one who has shown that meanings—propositions; that clauses—can be viewed as hypostacizations of our linguistic practices has actually deflated meaning, as opposed to having merely deflated *meanings*. I think that providing such an account is an important step towards deflating meaning; however, in order to truly deflate *meaning*, it seems to me that we must show that we can provide an account of the explanatory role of meaning in terms of the linguistic role that the meaning predicate has in our language. The reason is that a deflationary account of meaning will really be a deflationary account of meaning *and* content, and it seems to be compatible with the fact that we have an account of our linguistic practice of ascribing meanings that content plays an

those locutions is not *just* to deny that truth and truth conditions play a role in meaning; rather, it involves the claim that we can account for the explanatory role of meaning without providing a reductive account of its nature. Field has not done this, and, to the extent that he has not, I think we can say that he has not yet provided a deflationary view of meaning or content.<sup>196</sup>

Of course one might argue that, although what I have just said is true, it is innocuous, because Field has not attempted to offer a deflationary view of meaning or content. What he has attempted to do, they might claim, is offer a deflationary view of the role of truth conditions in meaning or content, and this is compatible with his offering a rather substantial view of those notions, so long as the account of meaning does not amount to a reconstruction, in other words, of what it is for an utterance or mental state to have the truth condition that it has.

There is good evidence that this is how Field conceives the project of adducing a deflationary view of meaning and content. As we have sketched it previously, Field proposes an account of meaning that would be more accurately characterized as an account of what it is for an expression, phrase or utterance to be meaningful to a speaker, where this will involve the provision of the conceptual role of the expression or sentence in that person's idiolect. Of course an account of what it is for an expression or sentence to be meaningful for a person will not truck with meanings, and it will have no use for truth conditions, however it is developed. Armed with such a thing, would Field have a complete deflationist account of meaning? I cannot see that he would. To be sure, he

---

explanatory role in psychology. Please note that I am not claiming that content plays an explanatory role in psychology but that the claim that it does is compatible with the claim that we can deflate talk of meanings.

would not have a substantive theory of meaning; instead, he would have a sort of replacement account. What he would lack, however, is an account of the role of meaning ascriptions: if there are no meanings, and if, when we appear to ascribe meanings to things, we do not, we must provide an account of the utility of the meaning predicate, we must say what meaning ascriptions *do*.

As I see it, the proper way to understand Field's replacement notion of meaning is not as an account of what meaning is, though he provides us with some reason for believing that this is not how he is conceiving of it (Cf. his claim, in 1994a, p. 254, that "some versions of the view are in a sense quite undeflationist about meanings."). Instead, I think that it should be viewed on analogy with his claim (Ibid., p. 250) that "for a person to call an utterance true in this pure disquotational sense is to say that it is true-as-he-understands-it." Field goes on to remark that, in saying this, he is not giving an account or definition of truth; rather, he is providing a sort of heuristic, one that will be helpful in his discussion of the various features of disquotational truth. My proposal is that we understand the deflationist to take claims about "what meaning is" in the same way—they are useful heuristics that will help the deflationist to motivate her account of meaning, i.e., according to which there is no (substantive) objective property of meaning, etc. If this is correct, then the deflationist about meaning should not attempt to give an account of what meaning is. Instead she should uncover the function of 'meaning talk', and she should explicate meaning on the model provided by deflationary accounts of reference and truth.

---

<sup>196</sup> Of course I do not mean to say that anything that he says is incompatible with the provision of a deflationary account of meaning.

As we do not know what the function of ‘meaning talk’ is, we will not take this facet of the discussion any further. If our arguments against the Horwichian position are sound, the deflationist about truth needs to adopt a deflationary account of meaning. However, as we have seen, it is not exactly clear what that amounts to. On one reading of what it is to deflate meaning, a deflationary account of meaning, like a deflationary account of truth or reference, would have to do more than refuse to countenance a truth condition-based notion of meaning: it would have to include an account of the function of the meaning predicate that would provide us with the materials necessary to provide a non-reductive analysis of meaning. On the face of it, this is not the option that Field takes (N.B., though his overall program is consistent with this approach). On another reading of what it is to deflate meaning, a deflationary account of meaning is just a non-truth-conditional account of meaning. On this reading, as previously mentioned, we do not deflate meaning, so much as we deflate the role of truth conditions in meaning.

It is clear that Field has taken the second option, and that, contrary to what one might have thought, he has not offered a truly deflationary account of meaning. Furthermore, because denying truth conditions a role in meaning entails denying that there are objective meaning properties, he has offered an eliminative account of meaning: on his view, we cannot talk of the meaning of an expression, phrase or sentence, and we cannot fill in the ‘...’ in the schema

$\sigma$  means that  $p$  iff ...

Instead, we must replace questions of what an expression, phrase or sentence means with questions—would anyone ask these?—of how this or that person interprets or understands an expression, phrase or sentence, or, maybe, how best to interpret or

understand an expression, phrase or sentence. Hence, this non-deflationary view of meaning is, in the Quinian spirit, eliminative about meaning and meaning facts.

In itself, this is not a criticism of the view, of course. However, we might wonder on what basis the Fieldian deflationist about truth puts forward her eliminativist view of meaning. Indeed, even if we judge this view to be a non-eliminative account of meaning,<sup>197</sup> we still must ask why we should accept it. This is a problem with eliminativist views,<sup>198</sup> generally.

Notice that this problem does not arise with respect to a genuinely deflationary view of meaning, just as it does not arise with respect to a deflationary view of truth: on either one, it is part of the view that the explanatory role of meaning (truth) is... (where '...' is filled in with some description of the function of the concept of meaning (truth)), and that if the explanatory role of meaning (truth) is ... then, having explained the linguistic function of the meaning (truth) predicate, we have said everything there is to say about the nature of meaning (truth). The only question that remains, on this sort of account, is whether or not the deflationist is correct about the explanatory role. If she is then her account—or, at the very least, one like it—must be accepted.<sup>199</sup>

Field has not provided any argument for the account of meaning that he favors, and, hence, he has not provided us with a reason for accepting it.<sup>200</sup> This is surprising, in light of its importance to his deflationary theory of truth, for unless we have a reason to

---

<sup>197</sup> Of course we cannot deny that it is eliminative about meaning facts. Quibbling over whether or not to classify it as eliminative *tout court* strikes me as both silly and pointless.

<sup>198</sup> Witness the trouble had by the eliminativists in the philosophy of mind.

<sup>199</sup> *Modulo* the assumption that, having captured the explanatory role of a notion, we have said everything of interest that there is to say about it.

<sup>200</sup> Of course if it really is compatible with deflationism about truth, and I believe that it is, then he has provided the deflationist about truth with a reason for wanting to accept it, but that is a different issue entirely.

accept an eliminativist view of meaning,<sup>201</sup> then, unless we have on hand a deflationist view of meaning, we cannot accept a thorough-going deflationist view of truth. Insofar as Field has not motivated the view of meaning he sketched, he has not motivated the thorough-going deflationism that he has advertised.

There are two responses that I anticipate, and that I would like to try to block. The first response is that I have misunderstood Field's mission: he was not trying to motivate an account of meaning. All that he was trying to do was to say what a deflationist about truth should say about meaning. What follows from this is that, even if what I have said is true, it is besides the point.

In response to this, I should point out that I am not accusing Field of having erred, although I do think that the account of meaning that he was advertising was on the wrong track. What I was trying to show was this: first, deflationists do not need to adopt the account of meaning that Horwich offers, because they can adopt a deflationist view of meaning (if one can be developed); and second, they cannot adopt the account of meaning that Field is offering without an argument to the effect that there are no objective meaning facts. I assume that the meaning eliminativist would not want to rest everything with Quine's (or Kripke's) arguments, which means that she has a real argumentative burden. I also suggested that, if one wants a thorough-going deflationist account of truth, there is good reason for trying to develop a deflationary account of meaning. Hence, among other things, what I have tried to show is that, even if Field's account of meaning would be adequate to the purposes of the deflationist, it does not come for free.

---

<sup>201</sup> Or a deflationist view, of course.

I should note that this point is very important in Field's case, for Field advocates *methodological deflationism*, which is the view that we should assume deflationism unless pushed from it by certain explanatory needs. If I am right about his account of meaning, and, especially, about its importance to his deflationary account of truth, then it is clear that we cannot adopt methodological deflationism unless we can provide some reason for denying the factivity of meaning claims, or unless we can develop a genuinely deflationary account of meaning.

The next response that might be made is that the Fieldian Deflationist can adopt this view of meaning, since it has been arduously defended by numerous philosophers over the course of the 20<sup>th</sup> century. Of course this would be true if Field were willing to accept any of the arguments that have been proffered, but, owing to strength of the criticisms that have been leveled against those arguments, I doubt very seriously that he would take the arguments on board; and, anyway, it is not clear that his view is compatible with the verificationist or neo-verificationist positions that at least some of those arguments support. Hence, I believe that if Methodological Deflationism is to be accepted, either we need a genuinely deflationary account of meaning, or we need an independent argument for the non-factivity of meaning. Until we have one or the other, we cannot be methodological deflationists.

### **Chapter 5: The Contingency of Sentential Truth Conditions**

In this chapter, I would like to discuss a problem that has been raised by many philosophers, but that, as far as I can tell, has never received a satisfactory answer by the deflationist, viz., the problem of the contingency of sentential truth conditions. The

problem of the contingency of sentential truth conditions (herein, CTC), is the problem posed by the putative fact that the sentences of natural language have their truth conditions *contingently*, meaning that if relevant facts (e.g., about our usage) had been different, they would have had different truth conditions than they actually do. In the discussion of this issue, it has been assumed by the substantivists that we should assume that non-propositional truth bearers<sup>202</sup> have their truth conditions contingently and we should be moved from that position only if the deflationist offers us an argument that suffices to convince us that it is false. I think that this assumption stacks the deck against the deflationist and, for that reason, I will not countenance it.

What I will assume is no more than must be granted by every person who has ever thought about the issue: we have the intuition that the sentences of our language could have had different truth conditions than they actually have,<sup>203</sup> and any theory of truth or truth conditions must explain this intuition or, if it cannot, there must be some way of explaining away the intuition that is consistent with the truth of that theory. In my argument, I will employ the version of deflationism that we developed earlier, i.e., the more liberal Horwichian version, though nothing turns on the employment of this version. As I will show, if we attempt to explain all of the facts about truth, for all potential truth bearers, by reference only to propositions of the forms listed in (S1)—

(S1)

Truth for sentences

(ES) “p” is true iff p

---

<sup>202</sup> For ease of discussion, we will speak of sentences.

<sup>203</sup> Although no one who discusses the issue is clear about this, I take it that what we are supposed to have an intuition about is an utterance: that if things had been different this utterance could have been of a type tokens of which have truth conditions different from the truth condition actually had by this utterance. I assume that what is meant is that the physical token that I utter could have been of a different semantic type.

Truth for propositions	(EP) The proposition that $p$ ( $= \langle p \rangle$ ) is true iff $p$
Truth for beliefs	(EB) A belief that $P$ is true iff $p$
Truth for utterances	(EU) An utterance of “ $p$ ” is true iff $p$

—and auxiliary resources, such the schema  $\text{EXP}(\langle S \rangle) = \langle S \rangle$ , then we cannot account for our intuition of the contingency of sentential truth conditions. Of course one rather straightforward way to account for the intuition is to accept the CTC, and to adopt an account of sentence truth that renders it contingent. If the deflationist gambits fail, I think that we should accept this alternative. But this alternative entails that sentence truth is non-deflationary. Hence, if the deflationist cannot explain away our intuition of the CTC, I think that we have good reason for seeking a more-than-deflationary account of sentence truth.

Let's start with the obvious. We all accept that, or, better, we all have the intuition that (77) is true.

(77) Although 'snow is white' is true iff snow is white, it might have been true iff tofu is poisonous.

More generally, we all have the intuition that the sentences of our language could have had different truth conditions than they actually have, the intuition of the contingency of sentential truth conditions, or CTC, for short. Suppose we were to explain the CTC, the putative fact about which we have our intuition. How would we do it? Here's one way of explaining it:

(a) The sentences of a natural language—English, to take a familiar example—have the truth conditions that they do because they mean what they do, and

(b) They mean what they do as a matter of convention,

by which I mean only that

(c) If things had been different, individual sentences of our language would have had different meanings than they actually have.

So, the conventionality of meaning entails that our sentences might have had different meanings than they do, while the dependence of truth conditions on meaning ensures that if a sentence had had a different meaning than it actually does, it might have had different truth conditions. Bringing these together, we find that the dependence of truth conditions on meaning, together with the conventionality of meaning explains CTC—sentences might have had different truth conditions than they in fact have.

All of this sounds obvious enough, but it creates serious problems for the deflationist. If we take CTC seriously, it follows that instances of (DS), “p” is true iff p, are contingent. The reason is this: given the dependence of truth conditions on meaning, together with the conventionality of meaning, we have it that, for the sentences of natural language, s, if (DS) holds for s then (CTC) holds for s

(CTC) Possibly  $\sim$ (“s” is true iff s),

and it follows from this that there is no sentence of English, s, such that  $\langle s \rangle$  explains  $\langle$ “s” is true $\rangle$ , which is another way of saying that, given (CTC),<sup>204</sup> the deflationist cannot explain sentence truth by reference to the instances of the disquotational schema. Hence, if we take the CTC seriously, then the deflationist cannot explain all of the facts about truth that need explaining by reference only to instances of (DS), together with uncontroversial auxiliary resources, from which it follows that she cannot explain truth.

Let me explain the problem. If we take seriously our intuition that the sentences of our language might have had different truth conditions than they actually have, then we will hold that instances of the schema (DS) are contingent. If we hold that instances of (DS) are contingent, we must be prepared to say on what their holding depends—what makes it the case that ‘snow is white’ is true iff snow is white, rather than iff tofu is poisonous, for example, given that it might have been true iff tofu is poisonous? One obvious explanation for why it is that ‘snow is white’ is true iff snow is white, rather than iff tofu is poisonous, given that ‘snow is white’ has its truth conditions merely contingently, is that it means that snow is white.<sup>205</sup> If we accept the contingency intuition, and we accept that a sentence has the truth conditions that it does because it has the meaning that it has, then we will explain the contingency of truth conditions in terms of the contingency of meaning. This comes out nicely if we make the assumption that for a sentence, *s*, to mean that *p* is for it to express the proposition that *p*: *ex hypothesi*, *s* has the truth condition that it does because it expresses the proposition that it expresses; if it had expressed a different proposition, it might have had a different truth condition. If we explain the contingency of truth conditions in these terms, then we must explain how sentences come to express the propositions that they do, and we must do so in a way that is compatible with the contingency of meaning, which means that it cannot be explained by reference to a schema such as (EXP), which schema, if successful, provides a

---

<sup>204</sup> So as to avoid confusion, I will use ‘(CTC)’ to refer to the schema, and ‘CTC’ to refer to the thing about which we have our intuition.

<sup>205</sup> For those, like Katz, who deny that sense determines reference, this will have to be weakened, but the weakening is irrelevant to the present point, which point depends only on the assumption that truth conditions depend on meaning, not that the meaning of a token exhaustively determines the truth condition of that token on that occasion of use. Hence, room is left for the claim that context plays a role in determining truth condition, as well.

deflationary account of what it is for a sentence to express a proposition. What we would need is a substantive theory of meaning—one that explains the contingent fact that our sentences express the propositions that they do, and, by extension, explains the CTC. However, when the deflationist allows that our explanation resources might include, aside from all instances of (DS), certain unproblematic auxiliary resources, she aims to exclude precisely what we are claiming she needs: a substantive account of meaning.

The problem is clear. We all have the CTC intuition, and we have an explanation (or, anyway, the beginnings of one) for the fact about which we have our intuition. However, this explanation invokes a substantial theory of meaning, which is something that the advocate of a thorough-going deflationism cannot do, consistent with her claim that all of the facts about truth can be explained by reference to the schemata mentioned previously, together with unproblematic auxiliary resources. What should the deflationist do, assuming that she obstinately holds on to her deflationist view?

### **5.1 Deflationist Responses**

There are three types of response open to the deflationist, the first of which is a challenge, the second of which is an objection, and the third of which is an attempt to accommodate our intuition without succumbing to our substantivist pressures. I cannot prove that these responses exhaust the range of possibilities, which means that it is always possible, at least in principle, that the deflationist could adduce some other consideration to counter our argument. However, as I hope will become obvious, the likelihood of there being some other response is unlikely. If we can show that none of the deflationist responses are successful then, since, as we have noted, the deflationist

about truth cannot explain the CTC (at least she cannot if she wishes to hold on to her deflationist view), we will have provided a reason for believing that thorough-going deflationism is fatally flawed, from which it would follow that the deflationist ought not to be a thorough-going deflationist. That said, I turn to the deflationist responses to the problem posed by our intuition of the CTC.

### 5.1.1 The Challenge

The challenge that might be leveled by the deflationist leans heavily on the qualification that the deflationist is required to explain only those facts that need explaining. Simply pointing to the aforementioned modal fact isn't showing that it needs explaining, the deflationist might complain. If the critic is to establish an explanatory burden, she must show the deflationist that the CTC (or, at any rate, our intuition of the CTC) needs explaining.

Rising to the occasion on behalf of the critic, we might point out that we have made plausible the claim that (CTC) derives from the constitutive link between meaning and truth, and, hence, that we have provided reason for believing that it is explanatorily non-basic. That is, we have shown such facts to admit of explanation, and, therefore, we have made it plausible, at least *prima facie*, that they should be explained. Though I believe this point to be adequate, we can make a stronger claim.

Recall that Horwich took the putative fact that instances of (DS) were conceptually basic to play an important role in his theory. Although he has not been as explicit as he could be about what role this fact is supposed to play, it is not difficult to ascertain its place in his scheme: if instances of (DS) are conceptually basic then, unless

we find some explanatory need that would be served by providing a (non-deflationary) *theory* of truth, there is simply no reason to expect a substantive account of truth to be forthcoming. Aside from the obvious argumentative advantage Horwich's claim about the conceptual basicness affords his account, it also allows him to circumvent the charge that his view is revisionist—after all, if instances of (DS) really are conceptually basic, then he is providing us with an account of our ordinary notion of truth.

The claim to be accounting for our ordinary concept of truth is important, then, because, if successful, it turns the tables on the substantivist, forcing her to provide a reason for thinking that the property of truth has features not reflected in our concept. Does the putative fact that instances of (DS) are a priori, obvious, and, hence, conceptually basic support a deflationist position? Well, it would, except for one small problem: instances of (CTC) are a priori, obvious, and conceptually basic, too.<sup>206</sup> If this is the case, then presumably any theory worth considering should have something to say about them, and, what is more important, any theory of truth purporting to capture our ordinary notion must have something to say about them. The deflationist either must explain them, or she must explain them away, but she cannot ignore them. Hence, whether we go with the first response, the second response, or both, it is clear that the challenge to say why the deflationist must explain (or explain away) our intuition of the CTC has been met.

### 5.1.2 The Objection

---

<sup>206</sup> What's my evidence for this claim? Well, I do not think that this is the sort of claim for which we should expect anything like substantial evidence. It is enough to point out that if we do have the intuition of the CTC, and, hence, if we do accept instances of (CTC), it is not likely to be the case that we accept it on the basis of empirical evidence (it is, after all, a modal claim).

The objection is a bit more interesting, for it assumes the need to explain the CTC. The objection turns on the point mentioned before, about the obviousness of the (putative) modal facts. If these facts are so obvious, the deflationist might figure, why not simply enrich the theory to include them, perhaps by adding to it all instances of the schema, (CTC)? We don't need to get mired in the details of the additions,<sup>207</sup> for what would result is obvious enough: one aspect of the account would imply material equivalences of the form "s" is true iff s, while another aspect of the account would imply the contingency of those equivalences.

If such additions could be made, the need to account for the modal facts would, indeed, be obviated. Would it follow that all of the facts that needed explaining would be explained? I do not think so, for now a new issue would arise. What would need explaining is why it is that the equivalences implied by the theory—the instances of (DS)—are *material* equivalences, rather than *necessary* equivalences. The obvious answer would involve pointing to the fact that for each sentence, s, for which an instance of (DS) holds, an instance of (CTC) holds. That is, the deflationist would explain why the equivalences entailed by the theory are material equivalences, rather than being necessary equivalences, by pointing out that they are contingent. The problem with this should be obvious, and there are two ways to bring it out. First, we might note that, since the additional schemata are merely stipulated to be part of the theory, they couldn't be enlisted in our explanation. Second, we might note that, if the deflationist attempts to

---

<sup>207</sup> For those who are interested, propositional forms of the following sort would most likely be added <(If possibly  $\sim P$  then possibly "P" is true and not P) & (If, possibly P, then, possibly "P" is not true and P)>.

explain the fact that the theory entails material, rather than necessary, equivalences, by reminding us that they only hold contingently, we would be in our rights to ask on what their holding depends—why are they contingent?—and we are back where we started. For these reasons, I do not think that the deflationist can explain our intuition by enriching her theory with more schemata.

### 5.1.3 Attempted Accommodations

The third response open to the Deflationist, which has been made by Hartry Field (1994a), involves attempting to accommodate our intuition. It is by far the most interesting and creative. In what follows, I will consider Field's discussion, pointing out where I think that it goes wrong. However, before we turn to Field's strategy, let's consider whether or not a deflationist account like Horwich's, combined, as it is, with his use theory of meaning, can accommodate the intuition more directly.<sup>208</sup>

#### 5.1.3.1 A Horwichian Accommodation

Recall that Horwich accepted the following schemata as central to our theories of truth, reference, and the 'true of' relation, or predicate reference.

(DS) "p" is true iff p

(R) "N" refers to x iff  $x = N$ <sup>209</sup>

---

The point of the antecedent of the first conjunct is to ensure against the false 'possibly '1 = 1' is true and  $\neg(1 = 1)$ .' Adjustments would have to be made for indexical expressions, of course, but the general flavor of the modal propositional forms that would be introduced is clear enough.

<sup>208</sup> Of course in doing this, I assume, for the sake of argument, that Horwich's mixed use theory of meaning/deflationist theory of truth works, even though, as we have previously shown, it does not.

<sup>209</sup> This must be fixed up a bit to handle things like vacuous names and the like, but we will not attempt to improve on it here. In particular, we will not replace (R) with Field's (1994a) reference schema, (R<sub>F</sub>) (R<sub>F</sub>) If b exists then "b" refers to b, though we could.

(PR) “F” is true of x iff X is F

As Horwich (1995, p. 359) writes “According to the deflationary point of view there is nothing more to our concepts of [truth], being true of and reference than is conveyed by our acceptance of these schemata.”

Suppose, as before, that a sentence has the truth condition it has because it has the meaning that it has, a predicate is true of something in virtue of having the meaning that it has and that a name refers to what it does in virtue of having the meaning that it has, and, as before, that sentences, predicates and names all have their meanings as a matter of convention, by which I mean, contingently.<sup>210</sup> As we know, it follows from this that, though a name actually refers to what it does, it might have referred to something other than what it refers to, and, though a predicate actually is true of something, if it had meant something different, it might have been true of something else, or, perhaps, of nothing at all.

Suppose, further, with Horwich, that meaning is analyzed as use. This seems to provide us with the means necessary to accommodate our intuitions after all. *Ex hypothesi*, any predicate, sentence, or name might have been used in a way that is different from how it is actually used, and, if it had been used differently, then, on Horwich’s account, it would have a different meaning from the meaning that it actually has. If this is the case, then, given the dependence of truth condition on meaning, we are guaranteed that instances of (PR) and the rest hold contingently. Hence, it looks as if

---

<sup>210</sup> I think that there is good evidence that names have meaning (Cf., for example, Katz, 1994, 1997, 1999). It is important to note, however, that whether or not they do is not at issue, for it is assumed by Horwich (Cf, esp. 1999) that they do.

Horwich's account, if only it could be made to work, could accommodate our intuition after all.

I do not doubt that if the account could be made to work, then, given the assumed dependence relation, it would accommodate—in fact, explain—our intuition. Though I have grave doubts about the plausibility of Horwich's account, in what follows, I will put those to the side. As we will see in a moment, if Horwich's account could explain the data, then, far from supporting his view, this would be a very *unwelcome* result, for it would undermine his claim to be able to explain everything by reference to the propositional forms that correspond to the instances of the disquotational schemata mentioned above.

We should note that Horwich (1998) might accept a version of the dependence claim mentioned previously, for he accepts the dictum, *no difference in extension unless there is a difference in meaning*. Of course this claim amounts to nothing more than (78) (78) Difference in extension  $\rightarrow$  Difference in meaning, and that is not the (CTC). In order to derive the contingency of the instances of (PR),<sup>211</sup> we have to add the assumption that a predicate could have meant a property not exemplified by all and only those things that actually fall under it. I think that this is a trivial consequence of Horwich's use theory, together with the fact that we could have used the predicates of our language in ways that differ from how we actually use them. If this is correct, then, in fact, Horwich can explain the CTC, after all.

The problem with this should be obvious: if Horwich can explain the CTC by reference to his theory of meaning, together with propositions corresponding to the

relevant instances of the schemata, then his theory is not deflationist about sentence truth, for it would explain certain facts about (non-propositional) truth (and truth conditions) by reference to a substantive theory of meaning, which, as we have said previously, is the hallmark of a more-than-deflationist account. Hence, we have further evidence that, consistent with his thorough-going deflationism, Horwich cannot combine a use theory of meaning with a deflationary theory of truth. This, in turn, points up the fact that the deflationists either need to find a linguistic role for ‘means’ or they need to adduce an argument or motivation for adopting a meaning-theoretic eliminativism.

### 5.1.3.2 Field’s Deflationary Strategy

Rather than denying that that the intuition of the CTC needs explaining, or attempting to enrich his theory in the hopes of explaining the contingency of sentential truth conditions, Field puts a new spin on our old modal intuitions, taking them to be special cases of more general intuitions about how to understand the assignment of truth conditions to sentences of foreign languages. According to Field (1994a, p. 276), the intuition that grounds the objection from the CTC can be captured by counterfactuals such as (79), and the “cash value” of (79) is (80).

(79) If we had used “snow is white” in certain very different ways it would have had the truth condition that grass is red.

(80) In considering counterfactual circumstances under which we used “Snow is white” in certain very different ways, it is reasonable to translate it in such a way that its truth conditions relative to the translation are that grass is red.

---

<sup>211</sup> I should note that this can be made to work for the other schemata as well. I owe this point to Matt McGrath, with whom I discussed it at a conference at Brown University in 1998.

Field is recasting talk of our own language with different properties as talk of a different language altogether—one used not by us, but by our counterfactual selves. The intuition that appears to be about the modal properties of our own language is recast as an intuition about how we would translate from this foreign language into our own.

It is important to note that, as we have characterized deflationist strategies, this is a characteristically deflationary move. Rather than attempting to explain the putative modal fact that sentences have their truth conditions only contingently, Field attempts to accommodate our intuition in terms which don't require adverting to substantial semantic properties of our sentences. Instead of explaining how it could be that, though 'snow is white' has the truth condition that snow be white, it might have had the truth condition that grass be green, Field tries to capture our intuition that is registered in (79), and he attempts to do so in terms that do not require positing that our sentences have their truth conditions only contingently. Hence, by reading the putative modal fact as a fact about how we would translate sentences, Field captures the phenomena without rendering contingent the assignment of truth conditions to sentences. In this way, he deflates the contingency of sentential truth conditions, i.e., by showing that we need not invoke a substantial theory of meaning to explain the putative data, which include the fact that we are inclined to accept sentences like (79), for the sentences that we understand.

The point of Field's deflationist maneuver, then, is to show that the deflationist can account for the contingency intuition without accepting the contingency of sentential truth conditions, and thus without having to invoking a substantive account of meaning. He argues as follows: if (80) captures the content of (79) then, since the deflationist accepts (80), she accepts (79), and if she accepts (79), she accepts, and in some sense can

accommodate, the intuition that we claim she cannot accommodate. Can the deflationist avail herself of this strategy and thereby avoid the problem that we have posed? There are two primary reasons why I think that she cannot.

Field's strategy might be objected to on the grounds that he fails to engage the substantivist, by refusing to explain the fact about which we have our intuition. On this view of things, the problem with Field's strategy is that it involves switching desiderata on us. The issue, we might argue, is not whether or not the deflationist could affirm the intuition, but whether or not she could explain the *fact* about which we have the intuition. The worry that she cannot explain that fact is not dispelled by Field's strategy. Of course, in assuming that it is a fact that the sentences of our language might have had different truth conditions than they actually have, we beg the question against the deflationist. The only *fact* to which we may legitimately point is the fact that people are inclined to accept that the instances of (CTC) register obvious and *a priori* facts about natural language. Hence, although tempting, we cannot avail ourselves of this strategy.

The claim that Field changes the subject, then, is borne of a failure to appreciate the deflationist strategy, and, therefore, we conclude that this objection is misguided. Field is in his rights to take as the primary datum the fact that people have the intuition registered in (79), and he is in his rights to take that intuition to be better understood as in (80). Hence, if (80) captures the "cash value" of (79), as Field tells us it does, then, with respect to our intuition about the contingency of the truth conditions of the sentence 'snow is white', it looks as if the deflationist strategy will be successful.

For the moment, then, let's assume that, with respect to the sentence 'snow is white', Field's deflationist strategy is successful. Still, one might worry that Field's

strategy, even if it could successfully explain our particular intuitions, i.e., our intuitions about particular instances of (CTC), is not adequate. We are supposing that Field's strategy is successful, which means that we are supposing that, for each sentence of our language or idiolect about which we have *the modal intuition*, i.e., the intuition registered in (79), Field could explain away our intuition by reference to something like (80). That is, he could accommodate our intuition without invoking a substantial theory of meaning. Hence, he could accommodate successfully each case that we could bring to him.

What about the generalized version of the intuition, i.e., our intuition about the modal properties of sentences of natural languages, generally, or even of every sentence of our own language or idiolect? Even if, using the aforementioned strategy, Field could run through every natural language sentence and provided a correlate to (80) for each one, he couldn't register our general intuition, since universal generalizations are not logical consequences of the set of all of their instances. For this reason, this strategy is unable to accommodate our *general* intuition, our intuition that might be captured by (81).

(81) The sentences of natural language (or even just of our idiolect) have whatever truth conditions they have contingently

Some might think that this is not a problem; after all, if he can handle every instance, who cares about (81)? I think that this response (N.B., versions of which are made to numerous problems about generalizations that have been pointed out to the deflationists) is gravely confused. The reason it is confused is that it is very plausible to hold—in fact, wildly implausible to deny—that we accept (79) and instances like it *because* we accept

the generalized version, and not the other way around. Hence, I conclude that Field's deflationist strategy cannot accommodate the general character of our intuition.

Although I think that this criticism is sufficient to show that the deflationist cannot accommodate everything that needs accommodating, let's not put too much weight on it. Instead, we'll stick to the instances of (CTC), to see whether or not Field's strategy will work on those. Suppose, as before, that we follow Field and take (80) to capture the cash value of (79). According to (80), in the envisaged circumstance, it would be reasonable to translate 'snow is white' so that it has the truth condition that grass is red. Let's assume this is the case, though, of course, I have said nothing about the circumstance we are supposed to envisage. We know, then, that, under certain conditions, it would be reasonable to translate the sentence as Field suggests, but we do not know why it would be reasonable to translate the sentence in that way. This raises a worry. The most obvious answer to the question of why it would be reasonable to translate the sentence in that way is:

- (i) in the imagined circumstances, 'snow is white' is being used to mean what 'grass is red' actually means;
- (ii) correct translations preserve meaning; and
- (iii) by 'reasonable' we mean 'correct as far as we can tell'

Although I think that this is the obvious strategy, it is not open to the deflationist, since (i) - (iii), and, in fact, any explanation that avails itself of objective correctness conditions, presuppose a substantive notion of meaning.

How else might the generalized deflationist explain the notion of reasonableness that she invokes? To be sure, Field suggests that she adduce pragmatic standards, but

even if those could be produced, it is clear that they would not be appropriate for the job at hand. There are various reasons for this. First, and most generally, if what we said previously about the general character is correct, then the whole translation strategy is suspect, since, in that case, our intuition covers the sentences of any actual or possible natural language, including actual languages that we do not understand and possible languages containing expressions that would resist translation into any actual language at all.

More importantly (as I see it), if the deflationist does not tie reasonableness to correctness—and she cannot—then we lose what we might call the *objective character* of our intuition. Here's what I mean: Suppose that we take (80) to capture the content of (79), but that we do not understand reasonableness to involve or presuppose either correctness or probable correctness (i.e., under any description). In that case, we could judge a translation to be both reasonable (as that is intended) and incorrect. However, the intuition Field aims to capture—that if we had used our words differently, 'snow is white' *would* have had a different truth condition—is incompatible with that; hence, (80) fails to capture the content of our intuition, from which it follows either that (80) does not capture the content of (79) or that (79) is not an adequate restatement of our intuition. Which option we take is really a matter of taste, for either results in the same conclusion: Field's gambit fails.

Let me turn to a response that one might think is available to the Fieldian deflationist,<sup>212</sup> though I doubt very seriously whether she would be inclined to take it.

---

<sup>212</sup> This response has been made by Matt McGrath, in the course of an email correspondence that we had. If there are any mistakes, they are most likely due to me, however, as I have not retained all of the emails that we wrote back and forth.

Given the argument that I have just developed, it might seem that if the deflationist could avail herself of an objective notion of synonymy—perhaps in terms of sameness of conceptual role—then she could answer my objection and avoid the problems we have tried to make for her. This would involve making use of schemata like (82) - (84), which

(82) There are counterfactual circumstances in which “p” as used by me has the same meaning as “q” as actually used by me.

(83) If “p” as used by me in a counterfactual circumstance has the same meaning as “q” as actually used by me, then in the counterfactual circumstance, “p” as used by me is truer iff q

(84) There are counterfactual circumstances in which “p” as used by me is true iff q. can be used for explaining—not just affirming, but explaining—facts like that captured in (79). This approach would obviate the need to explain, in a suitably un-meaning-theoretic way, why it would be *reasonable* to so-translate the sentence, and, hence, although it does involve positing objective synonymy relations, it might look as though it will allow the deflationist to avoid the problems we have posed.

I do not think that even this move will help the deflationist, however, and, thus, I do not think that the deflationist can avoid the problem posed by the putative contingency of sentential truth conditions, which problem, as previously noted, stems from the conceptual connection between truth and meaning. Aside from the fact that this solution posits objective synonymy relations, which, as Field (1986) notes, the deflationist can, but should not, countenance, the problem for this approach is the same as one of the problems mentioned earlier: our modal intuitions allow that the expressions of our language could have been such as to have meanings that are not synonymous with those

of the words of our, or any other, actual language. If things had been different—if, for example, we could conceive properties which we cannot actually conceive—then at least some of our words could have had meanings that resisted translation into even an extension of our actual language. Hence, contrary to what Field claims, using a translation strategy, we cannot capture the intuition registered in (79), which means that Field's attempt to provide a deflationist account of our intuition of the contingency of sentential truth conditions fails.<sup>213</sup>

If what we have argued is correct, then the three most plausible replies available to the deflationist will not allow her to avoid the problem created by our intuition of CTC; hence, short of denying that she has the intuition of CTC—a move that, to my knowledge, has not yet been tried—the advocate of a thorough-going deflationism must admit that her account of truth is not able to explain our intuition of the contingency of sentential truth conditions, and that, for this reason, she is not able to explain away our intuition of the contingency of sentential truth conditions.

If our argument for the claim that she cannot simply ignore the intuition is correct, it follows that a thorough-going deflationist account of truth is unable to deliver an account of truth (and meaning) that does everything that we have come to expect of such an account. If the deflationist cannot accommodate these intuitions, and if, as we have claimed, these intuitions must be accommodated, then, given that we can accommodate them on the hypothesis that, in fact, there is a dependence relation between the meaning of a sentence and its truth conditions, it follows that, at least tentatively, we must accept that there is dependence relation that exists between the meaning of a

---

<sup>213</sup> I would like to thank Matt McGrath and James Woodbridge for helpful discussions of many of the issues

sentence and its truth condition. If this is the case, then, since the deflationist holds that all of the facts about truth that need explaining can be explained by reference to the instance of (DS), together with uncontroversial auxiliary facts and theories (of which the relevant theory of meaning, whatever that turns out to look like, would not be one), it follows that we cannot accept a thorough-going deflationism. Of course, we cannot detach this last claim unless we can establish that there is no way for the deflationist to accommodate the relevant modal intuition, and I am not sure that this can be done. Nevertheless, the burden is not on us to show it is impossible for the deflationist to accommodate the data; given our argument, the burden is on the deflationist to show that, appearances to the contrary notwithstanding, there is a way for her to accommodate it.

Before, we close, we should note one final gambit that could be made by the deflationist, one that is in the spirit of much of what Field says, although, at least to my knowledge, he has not tried it. Suppose that we are correct and that the deflationist cannot explain our intuition of CTC. So what? If, as the deflationist claims, the explanatory role of truth is exhausted by the logical function of the truth predicate, then, so long as successful performance of that function does not depend on the contingency claim, why cannot the deflationist simply say that, though we have the intuition registered in (79), the notion of truth that we need—essentially, a disquotational notion—does not have that characteristic. If this notion is not our ordinary notion, so much the worse for our ordinary notion.

As it stands, I do not think that this response is really open to the deflationist, at least it is not open to the *methodological deflationist*, who believes that we should

---

raised and points made in this section.

assume deflationism (actually, disquotationalism), and enrich our account of truth only as that becomes necessary. The reason is that this would require assuming to be true a view that is freely acknowledged to be counter-intuitive. In general, we should not assume as a working hypothesis a view that has counter-intuitive results unless there is an argument that provides us with a good reason for doing so.

Fortunately for the deflationist, Field has an argument which, if it goes through, establishes that there is a reason for preferring a use-independent notion of truth to a use-dependent one, even if the former is less intuitively motivated than the latter.<sup>214</sup> Hence, although Field's argument was not developed for the purposes of responding to our challenge, it can be used to (attempt to) establish that the explanatory role of truth requires that we have on hand a use-independent notion of truth, and if it can do that, then the deflationist may be in her rights to ignore our intuition of the CTC, as well as everything that we have claimed follows from it. Here is Field (1994a, p. 265-266):

Consider some theory about the physical world, formulated with finitely many separate axioms and finitely many axiom schemas (each schema having infinitely many axioms as instances). For an example of such a theory, one can take a typical first order version of the Euclidean theory of space (which is not finitely axiomatizable)... Suppose that Euclidean geometry is true, and that we want to express its contingency by saying that the axioms together might have been false. Surely what we want to say wasn't simply that speakers might have used their words in such a way that the axioms weren't true, it is that space might have

---

<sup>214</sup> I should note that the claim that use-independent view is less intuitively motivated than a use-dependent view is not clearly true, given the fact that a propositionalist view is both use-independent and intuitively well-supported. Field is not advocating a propositionalist view, however, so we can safely ignore this fact.

differed so as to make the axioms *as we understand them* not true. A use-independent notion is precisely what we require. (italics original)

If Field is right then the deflationist has good reason for thinking that, even if she cannot accommodate the intuition lately registered, if truth is to play the explanatory role that gives it its point, then so much the worse for our intuition. By extension, if our ordinary notion of truth (if there is such a thing) is as of a use-dependent notion, then so much the worse for our ordinary notion of truth.

Before we consider Field's example, we should note that Field grants that the propositionalist also has on hand a use-independent notion of truth, since the truth or falsity of a proposition, as well as its truth condition, is independent of us and our practices, and, therefore, that the consideration adduced above supports either a disquotationalist view or a propositionalist view. When we combine this concession with Soames' claim—that our ordinary notion of truth takes propositions as the primary bearers of truth—it follows that we can accuse Field of attempting to replace our ordinary notion with his sententialist notion. In what follows, we will leave this concern to the side and consider whether or not Field's argument goes through.

According to Field, we need a use-independent notion of truth rather than a use-dependent notion of truth because, given the latter, we cannot use a sentence like (85)

(85) The axioms of Euclidean geometry together might have been false

to express the contingency of Euclidean geometry. When we utter such a sentence, Field reminds us, what we mean is that space itself might have differed so as to render the axioms, taken as we understand them, false; what we do not mean is that our use of the language might have differed so as to render the axioms (trivially) false. This is

important, because the latter is compatible with the necessity of space being as it is, and, therefore, couldn't be what we mean to say when we utter (85).

There are two important things to note. First, Field says "Surely what we wanted to *say* wasn't simply that speakers might have used their words in such a way that the axioms weren't true..." (italic added). I am unclear what he means by 'say' here.

Without pressing too hard on this, I think that it would be safe to say that what he means is that, in uttering the sentence, (86), what we want to *convey or communicate*

(86) The axioms taken together might have been false,

is that space itself might have differed so as to render these axioms as we understand them false.<sup>215</sup>

As Field is well aware, if truth were use-dependent then (86), if true, would be true because either

- (a) We might have used our words differently, or
- (b) Space itself might have differed.

This leads to our second point, which is that Field's suggestion, that if truth were use-dependent, then what we convey or communicate (i.e., *say*) in uttering (86) is "simply that speakers might have used their words in such a way that the axioms weren't true," is false. What we convey or communicate is the disjunction of (a) and (b). Hence, as Field is well aware,<sup>216</sup> in uttering (86), and given a substantive (e.g., correspondence) theory of truth, I could still convey the belief that space itself might have differed. The reason is

---

<sup>215</sup> Note, for example, that what he could not mean is *express the proposition or mean*, at least not consistent with his renunciation of objective correctness conditions. indeed, there is some basis for my reading. In Field (1986, p. 59), he says "The fact remains however that the belief that we are trying to convey [in uttering (8)] does not involve correspondence truth."

<sup>216</sup> See Field, 1986, p. 58, where he makes the same point.

that the relevant facts (i.e., captured in (a)) might be common knowledge, and *modulo* those commonly known facts, an utterance of (86) conveys (b). I conclude, therefore, that Field's claim that we require a use-independent notion of truth in order for truth to play its logical function is somewhat exaggerated.

If we are right that Field's claim is exaggerated, then, unless the deflationist can either explain or explain away our intuition of the CTC, I do not see that methodological deflationism can be adopted. Furthermore, unless the deflationist can either explain or explain away our intuition of the CTC, I do not see that she can deny the conceptual connection between meaning and truth. However, if, as we argued, the deflationist cannot explain the connection between meaning and truth, i.e., if she cannot explain the meaning to truth conditionals, then it looks as though she must take them to be explanatorily basic, which is a result that would not sit well with any deflationist. We conclude, therefore, that the deflationist must explain away our intuition of the CTC, or, at the very least, explain it without taking seriously the contingency of sentential truth conditions. It has been the burden of this section to show that this task is more difficult to accommodate than might have been realized.

Furthermore, if, as we have argued, (CTC) is conceptually basic, then, unless the Property deflationist can explain it away, she cannot lay claim to our ordinary concept of truth. As discussed previously, this has unfortunate results for her argumentative strategy, for if her account yields counter-intuitive predictions, then she cannot turn the tables on the substantivist. Recall that the import of the property deflationist's claim to be capturing our ordinary notion of truth was that, if correct, it entailed that the substantivist had to adduce reasons for adducing a more-than-deflationary account of the

property of truth. However, if (CTC) is conceptually basic and if the deflationist can neither explain it nor explain it away, then she must provide the substantivist with a reason for adopting the deflationist position, rather than the other way around. We can see the import of our argument from the CTC: it furnishes a reason for expecting a more-than-deflationary account of sentence truth that is consistent with the deflationist's claim that the explanatory role of truth is exhausted by the logical function of the truth predicate.

## 5.2 Where Do We Go From Here?

At this point, it would be reasonable to assume that, if our arguments have been convincing, it is appropriate to turn back to substantivist accounts, to see what they have to offer. I believe that many philosophers, finding some problem with deflationism, have been too quick to affirm its obverse, substantivism about truth for all truth bearers, or *generalized substantivism*. Generalized substantivism isn't supported by our arguments because, if successful, our arguments establish only that we must give a more-than-deflationary account of truth for putative truth bearers—sentences, for example—whose truth conditions depend on their meaning or content. There are some truth bearers whose truth conditions do not depend on their meaning or content, however—propositions.<sup>217</sup> Propositional truth conditions do not depend on the meaning of propositions because, as we should conceive it, rather than having meaning or content, propositions *are* meanings or contents. Hence, while our arguments might suggest that we should not be thorough-

---

<sup>217</sup> In Katzian parlance, *external propositions*.

going deflationists, they do not suggest that we should accept generalized substantivism, either.<sup>218</sup>

What our arguments show is that the deflationist has a problem accounting for the connection between meaning and truth. Suppose, then, that we were to decide that the truth or falsity of a non-propositional truth bearer depended, in part, on its meaning, and that, for true utterances and beliefs, at least part of our explanation for why they are true (or false) would have us mention the fact that they have the particular meaning or content that they have. What would our account of truth look like?

First, we would note that to talk about *the* meaning of a sentence or *the* content of a mental state is to talk about the proposition it expresses; and that to talk about a theory of meaning or content is to talk about an account that explains what it is for an utterance to express the proposition that it does, or what it is for a mental state to have the content that it has.<sup>219</sup> The claim that propositions are meanings, together with the assumption that which truth condition a non-propositional potential truth bearer has depends on its meaning, entails that the assignment of truth conditions to an utterance, *u*, depends on which proposition *u* expresses. Presented at this level of abstraction, this picture is compatible with nearly every account of meaning ever developed, and is, therefore, platitudinous.

---

<sup>218</sup> At this point, it is important for me to point out, or, at least, to remind the reader, that the deflationists have given us no good reason for being either deflationist or skeptical about meanings. The reason, as noted earlier, is that that, unless the deflationists rest on a Quinian argument, which they should not, they have not shown that there could not be (inflationist) meanings, and they have not provided a deflationist reading of meaning talk, where a deflationist reading of meaning talk explains the *function* of the 'means that' locution without adverting to substantial properties. Hence, I do not even think that we have motivated a non-inflationist position when we have shown that we can account for talk of *meanings* pleonastically. What would remain to be described is why we have the 'means that' locution, i.e., what function it plays in our language.

However, what our arguments show is that our theory of meaning must assign propositions to sentences in such a way that it is a contingent fact that a sentence or utterance has the meaning that it has. Hence, whatever theory of meaning we adopt, it will either have to show to be contingent the fact that a sentence expresses the proposition that it does, or, at the very least, it will have to be consistent with an explanation for why we accept the contingency of meaning. Armed with such an account of meaning, we could explain what it is for an utterance to be true in two steps. Step one would involve explaining what it is for a sentence to express the proposition that it does, which explanation would be provided by our theory of meaning.<sup>220</sup> Step two would involve explaining what it is for a proposition to be true or false. We could then say that a sentence or utterance is true (or false) iff it expresses a true (or false) proposition.

The obvious next question is: What account of truth do we offer for propositions?<sup>221</sup> The arguments that we have adduced so far provide us with reason for being skeptical about deflationist accounts of non-propositional truth, but they do not provide us with reasons for being skeptical about deflationist accounts of propositional truth. Hence, given everything that we have said so far, we have no reason to deny a deflationist account of propositional truth.

This suggests the possibility of adopting a hybrid view of truth, one which combines a deflationary account of propositional truth with more-than-deflationary

---

<sup>219</sup> Note that, for the Katzian, this would be only part of a semantic theory, in particular, the part that dealt with bridge laws connecting internal to external propositions.

<sup>220</sup> I leave aside questions about the assignment of content to beliefs. Although I believe that the story would be similar in many ways to the one I am telling about sentences and their meanings, I will not develop it here.

<sup>221</sup> Some might think that the obvious next question is What does our theory of meaning look like? I do not think that this is *our* obvious next question, however, though it would be were I attempting to sell the hybrid view that I am sketching.

accounts of truth for non-propositional truth bearers, like sentences, utterances, beliefs, and the like.<sup>222</sup>

In a moment, I will turn to deflationist accounts of propositional truth. However, for now, we should note where we are. If our arguments were successful, and if we are inclined towards a deflationist account of propositional truth, then we will advocate a hybrid view of truth: substantivist for non-propositional truth bearers and deflationist for propositions. Of course someone might object to this mixed account, thinking it unparsimonious to employ one sort of account of truth for beliefs, sentences and utterances and another for propositions. Although I do not think that this is a particularly compelling criticism, I'll address it briefly. There's a very plausible principle of metaphysics, which I call the Differential Property Principle, or DPP. According to DPP, if a thing has a differential modal or dispositional property—a modal or dispositional property not shared by all other things—then there must be some non-dispositional, actual property of the thing in terms of which we can explain why it has this dispositional or modal property while other things do not.<sup>223</sup> Sentences appear to have the modal property of possessing their truth conditions only contingently, whereas propositions do not seem to have that property.<sup>224</sup> Assuming that sentences and utterances do have that property, whereas propositions do not, then, by the DPP, there must be some properties

---

<sup>222</sup> Of course in saying this, I am not claiming that I have a theory of meaning or content; only that our argument so far has led us to view such a position as potentially attractive.

<sup>223</sup> Something like the DPP has always seemed to me to underwrite Armstrong's famous criticism of Rylean behaviorism. Recall that Armstrong criticized Ryle's account of dispositions, preferring an identification of dispositional states with categorical bases of behavior.

<sup>224</sup> Some might object to the claim that beliefs have the property of having their truth conditions only contingently. I think that such a claim is based on a mistake, however, and that addressing it has important consequences for our understanding of the contingency of sentential truth conditions. The response to the objection turns on what exactly we assign truth to when we assign truth to a belief. If what we assign truth

had by sentences and utterances, and not had by propositions, that explains this. The likely candidates are meaning-determining properties of some sort or other. Hence, we have a principled motivation for providing one type of account of truth for sentences and utterances and another one for propositions—in fact, we have a principled motivation for providing accounts of truth for non-propositional truth bearers that explain their differential modal properties, which is exactly what this mixed account would do.

### 5.3 Propositional Deflationism

My purpose in this last section was not to advocate for a hybrid view of truth, although I believe that such a view might be very attractive, in light of the considerations adduced in this dissertation. Rather, I wanted to note that there were limitations to what our arguments, if successful, showed, and I wanted to sketch, albeit very briefly, the sort of hybrid view of truth that I thought lent itself to the picture that was emerging. The hybrid view is, or, at any rate, could be,<sup>225</sup> substantive about meaning, and, therefore, about what it is for a non-propositional entity to have the truth condition that it has. However, it does not follow from this that the view is substantivist about truth.

When considerations are restricted to the primary bearers of truth—propositions, i.e., what is said or thought when we say or think what we do—the account of truth is thoroughly deflationist, for it is consistent with the claim that there is no non-trivial general characterization of what it is for a proposition to be true or false. As we would describe the hybrid view, it acknowledges that truth may be a trivial property, though it

---

to is the psychological correlate to an utterance—a dated belief state—as I think that it is, then the claim that beliefs have their truth conditions contingently is, or should be, as plausible as the claim that utterances do.

denies that the fact that certain utterances have truth conditions, while others do not, marks a potentially non-trivial fact about those utterances.<sup>226</sup> The hybrid view is deflationist about propositional truth without being deflationist about non-propositional truth.<sup>227</sup>

At the beginning of this dissertation, I made a distinction between propositional deflationism and semantic deflationism. The distinction is now being put to work. To be deflationist about propositional truth is to hold that there is nothing more to propositional truth than is captured by the instances of the equivalence schema, (EP). To accept that is (EP) The proposition that  $p$  ( $= \langle p \rangle$ ) is true iff  $p$ .

not yet to accept any claim whatsoever in the philosophy of semantics, properly so called—propositional deflationism is not a semantic doctrine. Rather, propositional deflationism is a view about the metaphysics of truth, and about nothing more. I believe that propositional deflationism is an obvious view, especially in light of what propositions are taken to be.<sup>228</sup> If propositions are thought of as contents—i.e., as what we say or think when we say or think what we do—then the question,

- Under what condition is the proposition *snow is white* true?,  
amounts to the question,

---

<sup>225</sup> I add this hedge because I believe that certain pleonastic accounts of meaning could be slotted in. They might be deflationist about meaning in certain respects, but they would not necessarily be eliminativist about meaning.

<sup>226</sup> The point generalizes to other non-propositional potential truth bearers, of course. We are using utterances as a sample type of non-propositional potential truth bearer.

<sup>227</sup> Of course, though it is compatible with deflationism about non-propositional truth, it is also at home with a substantivist view of non-propositionalist truth.

<sup>228</sup> Note that this claim is not the same as the claim made by Jackson, Oppy, and Smith (1994), that if we take propositions as the primary truth bearers then the triviality of truth derives from the concept of proposition that we are working with, rather than from something about truth, *per se*.

- Under what condition is what you say or think true, when you say or *think snow is white*?

There can be no more trivial question. What you say or think, when you say or think *snow is white*, is true iff snow is white. What you say or think, when you say or think that London is beautiful, is true iff London is beautiful; furthermore, if what you say or think is true, when you say or think that London is beautiful, then it is true *because* London is beautiful. The same holds for all such cases.<sup>229</sup> There is no explanatory gain to be had from an attempt to unify all of these individual explanations, because they are already as simple and obvious as can be. More importantly, that this is so is, in an important sense, independent of the question, What makes it the case that in the circumstances—referring to a circumstance in which one says or thinks that London is beautiful—what one says or thinks, in uttering or cognizing what one does, is that London is beautiful, rather than something else or nothing at all? To say that it is independent of this question is to say that it is compatible with either a deflationist or a substantivist answer to it.

I think that there is considerable intuitive support for the claim that propositional truth is trivial: if we have reason for ascribing truth primarily to what we say or think, as (e.g.) Soames and others think we do, then we have reason for thinking that, restricted to its primary application, truth is not the sort of thing that admits of a robust or substantial explanation.<sup>230</sup> Hence, I think that Field is right: if we restrict our attention to interpreted utterances and we ask, As interpreted, under what condition is this utterance

---

<sup>229</sup> I ignore issues having to do with vagueness, indexicality and the like, although I recognize that they might arise at this point.

true?, our answer will be utterly trivial. What I have worried about occurs before we get to that point, however.

What is it to deflate propositional truth? In chapter 0 of this dissertation, I made a distinction between metaphysical deflationism and metasemantical deflationism. Whereas the latter is a view about the constellation of semantic concepts, the former is a view about truth and other “metaphysical” concepts, like property exemplification. As I described it, the traditional debate about the nature of truth, the one that takes off from Aristotle’s famous dictum and Pilate’s alethic curiosity, is, ostensibly, concerned with the former and not with the latter, though, of course, an answer to the latter entails an answer to the former. Propositional deflationism engages in the former debate, rather than the latter debate. By deflating propositional truth, we deflate the central ontological notions that made early correspondence views appear suspect, i.e., the notions of facthood and property exemplification. Hence, in many ways, propositional deflationism answers the traditional question, as well as the pre-Tarskian worry about the evils of an inflationist metaphysics.

One of the central worries about correspondence truth regarded the possibility of comparing propositions with facts, in order to determine their truth value, and, more generally, it regarded the dubious status of these facts to which all and only true propositions were said to correspond. Propositional deflationism is deflationist about those entities, for, on this view, though we can talk about propositions being made true by facts, such talk will be seen as a trivial consequence of our acceptance of (EP), together with the effects of certain harmless nominalizing transformations, such as would

---

<sup>230</sup> Of course this is compatible with the claim that what makes a non-propositional truth bearer have the

take us from ‘snow is white’ to ‘it is a fact that snow is white’ to ‘the fact that snow is white.’ On this view, facts are mere hypostatizations. Propositional deflationism is deflationist about truth and about the ontological extravagances that have been posited in the explanation thereof, but is not deflationist about anything else. In this way, it captures the intuitive aspect of deflationism, while rejecting the counter-intuitive aspect of substantivism, to wit, the ontological and ideological excrement of the traditional correspondence theory, though it is not committed to a deflationist semantics.

There is one final issue that I wish to address in this section. As we know, according to Horwich (1997, pp. 95-96),

The basic thesis of deflationism... is that the disquotational schema

“p” is true iff p

is *conceptually fundamental*. By this I mean that we accept its instances in the absence of supporting argument: more specifically, without deriving them from any reductive premises of the form

$x$  is true =  $x$  is F

which characterizes traditional (‘inflationary’) accounts of truth, such as the correspondence theory, the coherence theory, the verificationist theory and the pragmatist theory.

Let’s assume that he is correct, that that is the fundamental thesis of deflationism. I want to show that Horwich’s claim about the “fundamental thesis of deflationism” can be captured by the propositional deflationist, even if she combines her propositional deflationism with a substantive theory of meaning and content. This is because the

---

truth condition that it has is non-trivial.

propositional deflationist can provide a deflationist explanation of our acceptance of the instances of (DS), where, you will recall, a deflationist explanation of (DS) explains our (DS) “p” is true iff p acceptance of its instances by reference to our grasp of the concept of truth, rather than by reference to our grasp of any reductive premises such as might be found in a substantivist account of truth. If I can show this, it will be an important result, because it will show that the claim that instances of (DS) are obvious, a priori, etc., which, *ex hypothesi*, constitute the basic thesis of deflationism, and which philosophers like Quine took to follow from the disquotational nature of truth, does not allow us to select between propositional deflationism and a thorough-going deflationism.

In what follows, I will show how a propositional deflationist might provide a deflationist explanation of (DS). To get started, we should mention that, in a footnote (1998b, p. 95, n. 1), Horwich notes that he takes ‘true’, as applied to utterances, to mean ‘expresses a true proposition.’<sup>231</sup> Hence, Horwich (1990, 1998b) explicitly accepts the schema, (87).

(87) “p” is true iff “p” expresses a true proposition.

Given the assumption that each utterance expresses at most one proposition, (87) is equivalent to (88)

(88) “p” is true iff what “p” expresses is true.

Given either (87) or (88), plus the claim, (89), and the schema that captures the heart of

(89) “p” expresses the proposition that p

---

<sup>231</sup> Although he does not note this, Horwich must make the assumption that each utterance expresses only one proposition. I assume that this is the case, in what follows.

propositional deflationism, (90), we can give a deflationist derivation of (DS).

(90) The proposition that p is true iff p.

Given (88) and (89), we get (91)

(91) “p” is true iff the proposition the p is true.

Given (91) and (90), we get (DS). Hence, we can derive (DS) from wholly

(DS) “p” is true iff p.

uncontroversial principles, whether or not we accept deflationary semantics.

It is easy to see that each of these premises would be acceptable to the propositional deflationist. (87) and (90) follow from propositionalism and propositional deflationism, respectively, while instances of (89) do no more than give the meaning of English sentences, and, hence, are known by those who know the meanings of English sentences.<sup>232</sup> If this is correct, then, contrary to what we might have thought, the triviality, obviousness, a prioricity or conceptual basicness of the instances of (DS) do not necessarily mark the divide between substantivist and deflationist views of non-propositional truth, since, as shown, it is compatible with each.

## **Chapter 6: Concluding Remarks**

The primary goals of this dissertation has been threefold: first, to get clear on what it means to accept a deflationary theory of truth; second, to investigate a class of problems that seem to plague thorough-going deflationist accounts, and to dispel myths about other alleged problems; and, third, to distinguish between deflationism about truth and deflationary semantics. As I have tried to show, a careful consideration of the

---

<sup>232</sup> It is crucial to recall that the claim about (DS) being conceptually basic is said to hold only for sentences one understands. Hence, (89) comes for free.

workings of a deflationist methodology provide us with reason for believing that philosophers are not yet in a position to adopt a deflationary attitude towards semantics. Furthermore, a consideration of the CTC, and the various ways in which the deflationist might try to accommodate it, has provided us with a reason for believing that the eliminativist views of meaning that have been proffered are unacceptable. Hence, if my arguments were good, I provided some reason for being skeptical about the prospects for a deflationist account of meaning.

Another very important issue emerged in this dissertation. My hope is that, in the course of investigating these issues, I have succeeded in illuminating—at least to some small extent—what it is to adopt a deflationist account of something. As I hope that I have shown, adopting a deflationist attitude towards X involves much more than simply denying the existence of the property of X, which is how some advocates *and* detractors of deflationism seem to understand it. It requires providing an explanation of the utility of the concept of X that does not require adverting to a property of X. That is, it requires attempting to explain the facts about a particular domain without providing reductive accounts of the key concepts employed in our talk about that domain. Hence, to take a relevant example, the central feature of a deflationist account of meaning is that a theory of meaning should not aim to provide reductive accounts of the locution *means that, has the content that, or expresses the proposition that*, but that it should account for what we do with those locutions by explaining their linguistic function, i.e., by explaining the utility of meaning and content ascriptions. Rather than explaining *meaning*, the deflationist explains our *talk* about meaning, which explanation does not require adverting to any underlying and substantive meaning properties. In this way, the

deflationist can accept that there are meaning facts, but, in so doing, she does not endorse the claim that such facts are susceptible of ontological reduction, or that they need be, in order for them to be rendered physicalistically acceptable, or just plain explanatorily kosher.

It is for this reason that I invoke the Carnapian distinction between internal and external questions.<sup>233</sup> The semantic deflationist deserves to be so called because of the form taken by her answers to the external questions about meaning and truth conditions, and because of the priority she gives to the internal questions: unlike the substantivist, who would claim that we can derive our answer to the internal questions—e.g., in virtue of what does the English word ‘snow’ mean *snow*?—from our answers to the external ones, the deflationist reverses the order of explanation, taking our answer to the external questions to be derivative on our answers to the internal ones.

If, as I have argued, this is the heart of deflationism, then the thorough-going deflationist must provide an account of the function of meaning ascriptions that is analogous to the one she provides for the truth predicate. However, as I noted, the deflationist about truth need not be a thorough-going deflationist. I claimed that there were in fact two separate projects that seem to have been fused (and, I would say, confused). The first is the traditional project of explaining the nature of truth, which project I have dubbed *propositional deflationism*. The second project is the project of deflating all of our semantic concepts, including meaning, which project seems to have developed in the wake of Quine’s anti-intensionalist arguments.. Success in the second

---

<sup>233</sup> As mentioned previously, the idea to employ the distinction I this context is due primarily to Huw Price, though I do not know whether or not our use of it is the same.

project entails success in the first, though, as I hope to have made clear, going in the other direction, this is not the case.<sup>234</sup>

I have argued that it might be difficult to succeed in the second project, given our intuition of the contingency of sentential truth conditions. I have not, however, argued that there would be a problem regarding success in the first project. In fact, think that to at least some extent, propositional deflationism is obviously correct, but that its correctness does not entail the correctness of metasemantical deflationism. I have also argued that propositional deflationism has the resources to explain the phenomenon taken by Horwich, Quine, Field and others to be fundamental to deflationism, i.e., the fact that, for all sentences a person understands, the disquotational schema will be accepted as an obvious, a priori truism. Hence, although I do not expect that someone like Quine, who is already a meaning skeptic, would be moved to adopt propositional deflationism, I have tried to show that it captures at least some of what the contemporary deflationist is after.

I believe that propositional deflationism is the heart of alethic deflationism, for it captures the sense in which there is nothing more to truth than can be captured by instances of the disquotational schema. However, the decision to accept it is not a metasemantical one. I understand the appeal of a thorough-going deflationism, as well as the appeal of a more explanatorily rich inflationist semantics, but I cannot see that issues regarding the explanatory role of truth can help us to decide which semantic view to favor. Hence, I think that questions about which we should prefer can neither be asked nor answered in the course of developing an account of truth.

---

<sup>234</sup> Of course for one who denies that there are propositions, the first project collapses into the second,

In closing, I should note that there are many issues that I have left entirely undiscussed, such as issues that arise in the course of handling the paradoxes, issues of vagueness, certain technical problems that plague the deflationist, and probably a host of other issues. Were I advocating seriously for a particular semantic picture, those issues would have to be considered in some detail. However, the issues that I have raised have not required a thorough consideration of those topics. Our interests were in getting clear on just what makes a theory of truth deflationary, i.e., on what seems to be central to the view; on why anyone would (or should) want to adopt such a theory; and on how to understand some important challenges that could be raised to claims of its adequacy. As I hope to have shown, what makes a theory deflationary has far less to do with the specific claims that follow from it, and far more to do with the philosophical methodology that it endorses. If this is correct, then the question of whether or not to adopt a deflationist account of truth cannot be asked until all of facts about truth that need explaining are clearly presented. It is my hope that in this dissertation I have made some progress towards that goal and have made even greater progress towards the goal of showing why it is of such great importance.

---

which explains why the projects have been (con)fused.

## References

- Alston, William (1996) *A Realist Conception of Truth*, Ithaca, NY: Cornell University Press.
- Ayer, A.J. (1935) "The Criterion of Truth," *Analysis*, 3.  
 (1936) *Language, Truth, and Logic*, New York: Dover Press.
- Blackburn, S. (1984) *Spreading the Word*, Oxford: Oxford University Press.
- Block, N. (1986) "Advertisement for a Semantics for Psychology," in P. French, T. Uehling, and H. Wettstein, 1986.
- Boghossian, P. (1990) "The Status of Content," *The Philosophical Review*, 69, 2, 157-84.
- Burge, T. (1979) "Individualism and the Mental," in French, P., T. Uehling, H. Wettstein (eds.) *Contemporary Perspectives in the Philosophy of Language*, Minneapolis: University of Minnesota Press, p. 73-121.
- Carnap, R. (1950) "Empiricism, Semantics, and Ontology," repr. In Carnap, R. (1956) *Meaning and Necessity*, 2<sup>nd</sup> ed., Chicago, Ill.: University of Chicago Press.
- Cartwright, R. (1987) "A Neglected Theory of Truth," in Cartwright, R. (1987) *Philosophical Essays*, Cambridge, MA: MIT Press.
- David, M., (1994) *Correspondence and Disquotation: An Essay on the Nature of Truth*, New York: Oxford University Press.
- Davidson, D. (1984) *Inquiries into Truth and Interpretation*, Oxford: Clarendon Press.
- Devitt, M. (1981) *Designation*, New York: Columbia University Press.  
 (1989) "Against Direct Reference," P. French, T. Uehling, and H. Wettstein, 1989.  
 (1990) *Realism and Truth*, 2d ed. rev., Oxford: Basil Blackwell.  
 (1996) *Coming to Our Senses*, Cambridge: Cambridge University Press.
- Dretske, F. (1981) *Knowledge and the Flow of Information*, Cambridge, Mass.: MIT Press.  
 (1990) "Does Meaning Matter?," in E. Villanueva, 1990.
- Evans, G. (1982) *Varieties of Reference*, ed. J. McDowell, Oxford: Clarendon Press.  
 (1985) *Collected Papers*, Oxford: Clarendon Press.
- Field, H. (1972) "Tarski's Theory of Truth," *Journal of Philosophy*, 69, 347-75.  
 (1973) "Theory Change and the Indeterminacy of Reference," *Journal of Philosophy*, 70, 462-81.  
 (1977) "Logic, Meaning, and Conceptual Role," *Journal of Philosophy*, 74, 379-409.  
 (1986) "The Deflationary Conception of Truth," in MacDonald and Wright (eds.) *Fact, Science, and Morality*, Oxford: Basil Blackwell.  
 (1994a) "Deflationary View of Meaning and Content," *Mind*, vol. 103, 249-285  
 (1994b) "Disquotational Truth and Factually Defective Discourse," *Philosophical Review*, vol. 103, 405-452.
- Fodor, J. (1987) *Psychosemantics; the Problem of Meaning in the Philosophy of Mind*, Cambridge, Mass.: MIT Press.  
 (1990) *A Theory of Content and Other Essays*, Cambridge, Mass.: MIT Press.  
 (1994) *The Elm and the Expert*, Cambridge, Mass.: MIT Press.

- Fodor, J. & E. LePore (1991) "Why Meaning (Probably) Isn't Conceptual Role," *Mind & Language*, 6, 4, 328-43.
- (1992) *Holism: A Shopper's Guide*, Oxford: Blackwell.
- Frege, G. (1952) *Translations from the Philosophical Writings of Gottlob Frege*, (ed.) P. Geach & M. Black, Oxford: Basil Blackwell.
- French, P., T. Uehling, H. Wettstein (eds.) (1979) *Contemporary Perspectives in the Philosophy of Language*, Minneapolis: University of Minnesota Press.
- (1980) *Midwest Studies in Philosophy*, vol. 5, *Studies in Epistemology*, Minneapolis: University of Minnesota Press.
- (1989) *Midwest Studies in Philosophy*, vol. 13, *Contemporary Perspectives I the Philosophy of Language II*, Minneapolis: University of Minnesota Press.
- Gibbard, A. (1993) "Meaning and Normativity," in Villanueva, ed., *Philosophical Issues 5: Truth and Rationality*, Atascadero, CA: Ridgeview, 95-115.
- Gibbons, J. (1999) "Truth in Action," ms.
- Grice, P. (1989) *Studies in the Way of Words*, Cambridge, Mass.: Harvard University Press.
- Gupta, A. (1994) "Minimalism," *Philosophical Perspectives*, 7, *Logic and Language*, 359-370.
- Harman, G. "Conceptual Role Semantics," *Notre Dame Journal of Formal Logic*, 23, 2, 242-56.
- Horwich, P. (1990) *Truth*, Oxford: Basil Blackwell.
- (1994) "What is it Like to be a Deflationary Theory of Meaning?," in E. Villanueva, 1994.
- (1995a) "Disquotation and Cause in the Theory of Reference," in E. Villanueva, 1995.
- (1995b) "Meaning, Use and Truth," *Mind*, Vol. 204, No 414.
- (1997) "Deflationary Truth and the Problem of Aboutness," in E. Villanueva, 1997.
- (1998a) *Meaning*, Oxford: Oxford University Press
- (1998b) *Truth*, 2<sup>nd</sup> Edition, Oxford: Oxford University Press..
- Jackson, F., G. Oppy and M. Smith (1994) "Minimalism and Truth Aptness," *Mind*, Vol. 103, No 411.
- Johnson, M. (1988) "The End of the Theory of Meaning," *Mind & Language*, 3, 28-42.
- Katz, J.J. (1977) "A Proper Theory of Names," *Philosophical Studies*, 31, 1-80.
- (1979) "The Neoclassical Theory of Reference," in P. French, T. Uehling, and H. Wettstein, 1979, 103-124.
- (1986) "Why Intensionalists Ought not be Fregeans," in E. LePore (ed.) *Truth and Interpretation: Perspectives in the Philosophy of Donald Davidson*, Oxford: Basil Blackwell.
- (1990) "Has the Description Theory of Names been Refuted?," in G. Boolos (ed.) *Meaning and Method: Essays in Honor of Hilary Putnam*, Cambridge: Cambridge University Press.
- (1994) "Names Without Bearers," *Philosophical Review*, 103, 1, 1 - 39.
- Kirkham, R. L. (1995) *Theories of Truth*, Cambridge, Mass.: MIT Press.
- Koslow, A. (1992) *A Structuralist Theory of Logic*, Cambridge: Cambridge University Press.

- Kripke, S. (1975) "Outline of a Theory of Truth," *Journal of Philosophy*, 72, 690-716.  
 (1977) "Speaker's Reference and Semantic Reference," in P. French, T. Uehling, and H. Wettstein, 1979.  
 (1980) *Naming and Necessity*, Oxford: Basil Blackwell.
- Leeds, S. (1974) "How to Think about Reference," *Journal of Philosophy*, LXX, 485-503.  
 (1978) "Theories of Reference and Truth," *Erkenntnis*, 13, 111-27.  
 (1995) "Truth, Correspondence, and Success," *Philosophical Studies*, 5-35.
- Lewis, D. (1973) *Counterfactuals*, Cambridge, MA: Harvard University Press.
- Loar, B. (1981) *Mind and Meaning*, Cambridge: Cambridge University Press.  
 (1982) "Conceptual Role and Truth Conditions," *Notre Dame Journal of Formal Logic*, 23, 272-83.  
 (1991) "Can We Explain Intentionality?," in Loewer and Rey, 1991.  
 (1992) "Self-Interpretation and Reference," in J. Tomberlin (ed.) *Philosophical Perspectives*, 8.  
 (1995) "Reference From the First Person Perspective," in E. Villanueva, (ed.), 1995.
- Loewer B. & G. Rey (eds.) (1991) *Meaning in Mind*, Oxford: basil Blackwell.
- McGinn, C. (1981) "The Mechanism of Reference," *Synthese*, 49, 157-186.  
 (1982) "The Structure of Content," in Woodfield, 1982.  
 (1983) *The Subjective View: Secondary Qualities and Indexical Thoughts*, Oxford: Clarendon Press.  
 (1984) *Mental Content*, Oxford: Basil Blackwell.
- McGrath, M. (1995) "Weak Deflationism," ms.
- Millikan, R. (1993) *White Queen Psychology and Other Essays for Alice*, Cambridge, Mass.: MIT Press.
- Neale, S. (1990) *Descriptions*, Cambridge, Mass.: MIT Press.
- Peacocke, C. (1981) "Demonstrative Thought and Psychological Explanation," *Synthese*, 49, 181-217.  
 (1989) "Perceptual Content," in J. Almog., J. Perry, & H. Wettstein, 1989.  
 (1992) *A Study of Concepts*, Cambridge Mass: MIT Press.
- Perry, J. (1990) "Individuals in Informational and Intentional Contexts," in E. Villanueva, 1990.  
 (1993) *The Problem of the Essential Indexical and Other Essays*, Oxford: Oxford University Press.
- Price, H. (1992) "Metaphysical Pluralism," *Journal of Philosophy*, 89, 387-409.  
 (1996) "Three Norms of Assertibility," ms.  
 (1997a) "Naturalism and the Fate of the M-Worlds," *Proceedings of the Aristotelian Society*, Supp. Vol. LXXI.  
 (1997b) "What Should a Deflationist About Truth Say About Meaning?," in E. Villanueva, 1997.
- Putnam. H. (1975) *Mind, Language, and Reality, Philosophical Papers*, vol. II, Cambridge: Cambridge University Press.  
 (1983) *Realism and Reason, Philosophical Papers*, vol. III, Cambridge: Cambridge University Press.  
 (1988) *Representation and Reality*, Cambridge, Mass.: MIT Press.

- Quine, W.v.O. (1953) *From a Logical Point of View*, Cambridge, Mass: Harvard University Press.  
 (1960) *Word and Object*, Cambridge, Mass.: MIT Press.  
 (1961) *Ontological Relativity*, New York: Columbia University Press.  
 (1991) *Pursuit of Truth*, Cambridge, Mass.: MIT Press.
- Ramsey, F. (1927) "Facts and Propositions," *Proceedings of the Aristotelian Society*, Vol. 7.  
 (1929) "Theories," repr. In his *Foundations*, ed. D. H. Mellor, London: Routledge & Kegan Paul, 1978.
- Recanati, F. (1989) "The Pragmatics of What is Said," *Mind & Language*, 4, 295-329.  
 (1993) *Direct Reference: From Language to Thought*, Oxford: Basil Blackwell.
- Richard, M. (1990) *Propositional Attitudes*, Cambridge: Cambridge University Press.  
 (1997) "Deflating Truth," in E. Villanueva, 1997.
- Russell, B. (1919) *Mysticism and Logic and Other Essays*, London: Longman's, Green and Co.  
 (1956) *Logic and Knowledge: Essays 1901-1950*, R.C. Marsh (ed.) London: George Allen & Unwin.
- Salmon, N. (1986) *Frege's Puzzle*, Cambridge, Mass: MIT Press.
- Schiffer, S. (1972) *Meaning*, Oxford University Press, New York and Oxford.  
 (1977) "Naming and Knowing," *Midwest Studies in Philosophy* 2, 28-41.  
 (1981a) "Indexicals and the Theory of Reference," *Synthese*, 49, 43-100.  
 (1981b) "Truth and The Theory of Content," manuscript.  
 (1987) *Remnants of Meaning*, Cambridge, Mass.: MIT Press.  
 (1988) "Symposium on Remnants of Meaning," *Mind & Language*, 3, 1-8; 53-63..  
 (1992) "Belief Ascriptions," manuscript.  
 (1993) "Actual-Language Relations," *Philosophical perspectives*, 7, 231-258.  
 (1994) "A Paradox of Meaning," *Nous*, 26, 279-325.  
 (1996) "Language-Created, Language-Independent Entities," *Philosophical Topics*, 6, 149-167.  
 (1996) "Meaning and Concepts," manuscript.
- Soames S. (1984) "What is a Theory of Truth?," *Journal of Philosophy*, 81, 411-29.  
 (1987a) "Direct Reference, Propositional Attitudes, and Semantic Content," *Philosophical Topics*, 15, 47-87.  
 (1987b) "Semantics and Semantic Competence," in S. Schiffer and S. Steele (eds.), *Thought and Language: Second Arizona Colloquium on Cognitive Science*, Tucson: University of Arizona Press.  
 (1997) "The Truth About Deflationism," in E. Villanueva, 1997.  
 (1998) *Understanding Truth*, New York, NY: Oxford University Press.
- Stampe, D. (1979) "Toward a Causal Theory of Linguistic Representation," in P. French, T. Uehling, and H. Wettstein, 1979.
- Strawson, P. (1950) "Truth," *Proceedings of the Aristotelian Society*, Vol. 24.  
 (1959) *Individuals*, New York: Methuen.
- Tarski, A. (1944) "The Semantic Conception of Truth," *Philosophical and Phenomenological Research*, 4, 341-76.  
 (1969) "Truth and Proof," *Scientific American*, 220, 63-77.

- (1983) *Logic, Semantics and Metamathematics*, 2<sup>nd</sup> ed., trans. J.H. Woodger, Indianapolis: Hackett.
- Villanueva, E. (ed.), (1990) *Information Semantics and Ontology*, Oxford: Oxford University Press.
- (1994) (ed.) *Philosophical Issues, 5, Truth and Rationality*, Atascadero CA: Ridgeview.
- (1995) (ed.) *Philosophical Issues, 6, Content*, Atascadero CA: Ridgeview.
- (1997) (ed.) *Philosophical Issues, 8, Truth*, Atascadero CA: Ridgeview.
- Wettstein, H. (1984) "How to Bridge the Gap Between Meaning and Reference," *Synthese*, 58, 63-84.
- (1986) "Has Semantics Rested on a Mistake?," *Journal of Philosophy*, 83, 185-209.
- Williams, M. (1986) "Do We (Epistemologists) Need a Theory of Truth?," *Philosophical Topics*, 14, 223-42.
- Woodfield, A. (ed.) (1982) *Thought and Object: Essays on Intentionality*, Oxford: Oxford University Press.
- Wright, C. (1987) *Realism, Meaning and Truth*, Oxford: Basil Blackwell.
- (1992) *Truth and Objectivity*, Cambridge, MA: Harvard University Press.