

# Suggested Answers to Philosophical Puzzles



# Suggested Answers to Philosophical Puzzles

By

Anguel S. Stefanov

**Cambridge  
Scholars  
Publishing**



Suggested Answers to Philosophical Puzzles

By Anguel S. Stefanov

This book first published 2022

Cambridge Scholars Publishing

Lady Stephenson Library, Newcastle upon Tyne, NE6 2PA, UK

British Library Cataloguing in Publication Data

A catalogue record for this book is available from the British Library

Copyright © 2022 by Anguel S. Stefanov

All rights for this book reserved. No part of this book may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without the prior permission of the copyright owner.

ISBN (10): 1-5275-8940-4

ISBN (13): 978-1-5275-8940-7

# CONTENTS

A Few Words About this Book .....	vi
Chapter 1 .....	1
Is the Gettier Problem a Stumbling Block for Epistemologists?	
Chapter 2 .....	13
Do Self-Referential Incoherent Theories Refute Themselves?	
Chapter 3 .....	24
What is Non-Classical Theory?	
Chapter 4 .....	31
What is Time Travel?	
Chapter 5 .....	46
Spacetime: Substantive or Relational?	
Chapter 6 .....	62
Is Spacetime an Emergent Entity?	
Chapter 7 .....	73
Does Big Bang Cosmology Resolve the First of Kant's Antinomies?	
Chapter 8 .....	95
Does the Anthropic Principle Explain the Appearance of Man in the Universe?	
Bibliography .....	108

## A FEW WORDS ABOUT THIS BOOK

In this book I undertake an attempt at suggesting answers to some philosophical puzzles. These are problems, which have still not ceased to be a bone of contention, bearing different reasons for the controversial theoretical assessments they have initiated.

Embarking on such an enterprise is a specific task and not an easy one for sure. It requires an attempt at adjusting and strengthening arguments for one of the already existing horns of a dilemma (for instance in chapter 5), finding a novel exit from a prolonged discussion (as is the case in chapter 1), defending unexpected answers (in chapters 3 and 4), and the like.

Some of the subjects of the chapters included in this book are thematically correlated, and others are not. To be “naturally” expected then, the suggested answers to the title questions of the thematically correlated chapters ought to stay in a conceptual harmony. Otherwise, discrepancies among them would certainly lower the plausibility of each of the suggested answers. This is a necessary requirement for the answers to the title questions of chapters 4-6. The reached conclusions of chapters 1, 2, 3, 7 and 8, are not thematically correlated.

# CHAPTER 1

## IS THE GETTIER PROBLEM A STUMBLING BLOCK FOR EPISTEMOLOGISTS?

### **The Gettier Problem**

It is widely conceded that the Gettier problem is a real epistemological problem. It appears to be a stumbling block for the epistemologists (or, to be more correct, mainly for the epistemologists within the analytic tradition) for one, but for a good reason. This reason is that the so-called Gettier cases demonstrate the incompleteness of the common presentation of knowledge to be standardly defined as *justified true belief*. A necessary amendment of this standard definition – the *JTB definition* for short – is thus required, so that it could deal with the exceptional Gettier cases, which are said to undermine the JTB definition. However, a good deal of attempts at suggested solutions for changes in the JTB definition of knowledge has been shown to be unsatisfactory in some aspect or another, and to this effect, to provoke continuous debates. This is the Gettier problem in a nutshell.

I shall try to defend a negative answer to the title question of this chapter, or to defend the claim that the Gettier problem ought not to be treated as a stumbling block for epistemologists, since *Gettier cases do not undermine the JTB definition of knowledge at a conceptual level*.

What Gettier (1963: 121-123) tried to demonstrate was *the thesis that the definition of propositional knowledge as justified true belief is not an adequate definition, since it is incomplete*. This means directly that there might be cases of justified true beliefs that do not present knowledge. Gettier himself adduced two examples of this sort, which are intuitively taken to be cases of lack of knowledge for an agent who is supposed to have a justified true belief.

The JTB definition of knowledge states that a subject S has knowledge of the proposition P, only if the following requirements are fulfilled:

- (a) P is true,
- (b) S believes that P, and
- (c) S is justified in believing that P.

According to the Gettier thesis this JTB definition of knowledge is incomplete, because it is not immune to exceptions, now bearing the name “Gettier cases”. They are of different types, but each of them is a case, in which all of the upper three requirements (a) to (c) are met, while it is still agreed that S does not know that P.

Let me refer to the first of the two original Gettier cases, which I’ll go back to later. The case is quite popular, so a brief retelling of it will suffice.

Smith and Jones have applied for a job, and Smith has heard the president of the company say that Jones will get the job. Meanwhile, Smith has managed to count the number of coins in Jones’ pocket and found that there are ten. So, having reliable evidence that Jones will get the job, and that Jones has ten coins in his pocket, Smith easily reaches the conclusion (marked by (e) in Gettier’s original paper):

- (e) The man who will get the job has ten coins in his pocket.

However, unbeknownst to Smith, he also has ten coins in his pocket, and he is the man who will get the job. On these factual grounds, Gettier insists that *Smith does not know that (e)*, although the proposition (e) is true, Smith believes that (e), and he is justified in so doing. According to the JTB definition, Smith ought to know that (e), but nevertheless (our intuition shows that) he does not know that (e).

The JTB definition fails to work in this case. Even if so, this fact does not yet set out a problem. If it is simply argued that some concrete definition of knowledge – the JTB definition – is incomplete, this may be avowed as



an epistemological *result* (no matter how significant), and not as an epistemological *problem*.

*A problem arises*, however, if the criticized JTB definition of knowledge has been officially and commonly accepted to be *the* definition of knowledge by the analytic philosophers before the publication of Gettier's paper in 1963. Even if this might not be the case,<sup>1</sup> a problem arises indeed, only if it has been accepted that the JTB definition of knowledge (even if not officially proclaimed) was the one, which was *tacitly embraced* by the epistemological community till the publication of Gettier's paper. Namely, this presupposition is widely shared and it is exactly this presupposition, which makes *the Gettier thesis emerge as a problem*.

### Attempts at a Solution

The persisting difficulty with the Gettier problem is that there has been no arguable solution for it to be commonly accepted by the epistemological community for sixty years since the publication of Gettier's challenging paper.

This being the case, how then can the Gettier problem be solved, or conceptually circumvented?

On the one hand, the JTB definition seemingly fails to be a general one. But on the other hand, its three requirements have never been evaluated to be epistemologically inadequate. This is why the first possibility directing the attempts at solving the Gettier problem, which occurred to philosophers, was the invention of some *fourth condition* for knowledge acquisition to be added to the requirements (a), (b), and (c), so that the new amended definition is not only a necessary, but also a sufficient condition for S to know that P.

---

<sup>1</sup> See in this connection the well-defended claim of Julien Dutant (2015) that the JTB definition *was not the traditional analysis of knowledge* till the publication of Gettier's paper in 1963.

After 1963 the justified true belief account of knowledge was seen to be defective and lost its exalted status; but even those convinced by Gettier that justification (along with truth) isn't *sufficient* for knowledge still mostly think it *necessary* and *nearly* sufficient for knowledge: the basic shape or contours of the concept of knowledge is given by justified true belief, even if a quasi-technical fillip or addendum ('the fourth condition') is needed to appease Gettier. (Plantinga 1990: 45, his italics)

Unfortunately, nobody has managed to provide such a universal "fourth condition" and to thus convert the JTB definition into one that is capable of resisting all types of Gettier cases. This way out of the problem has now been almost relinquished, since it has turned out that a "quasi-technical fillip" can hardly supply a plausible result.

Requirement (a) in the JTB definition cannot be altered, so the attempts at solving the Gettier problem were focused mainly on the elaboration of (c), and even of (b).<sup>2</sup> In the first original Gettier case, for instance, Smith believes that (e), due to the false presupposition that Jones will get the job. To this effect, the so-called "no false lemma" rule to the JTB analysis of knowledge was offered, expressing the principal claim that knowledge cannot arise from false premises. Plausible as it may seem, this stipulation proved to secure no general solution, because in some types of Gettier cases the knowing agent does not make conclusions from false premises, but she still fails to know.

Much hope was cherished for another solution, offered by the so-called "anti-luck epistemologists". In many Gettier cases the knowing agent S fails to know that P, in spite of the fact that S is justified in believing that P, and P is true, simply because *epistemic luck* has helped S to think that she knows that P. As a convincing example, I'll retell a short Gettier case, known as "the sheep-in-the-meadow case":<sup>3</sup>

---

<sup>2</sup> As B. Meyers-Schulz and E. Schwitzgebel (2013: 371) have argued for example, "we think there are cases in which it is intuitively plausible that a subject knows some proposition *P* without – or at least without determinately – believing that *P*".

<sup>3</sup> It is originally suggested by R. Chisholm, and I'm referring here to J. F. Rosenberg (2000: 30).

It's a bright sunny day and Smith is driving a car along a road, closely passing a meadow. He clearly sees a sheep in the meadow, so he has a justified true belief that there's a sheep in the meadow. What Smith really saw was a white rock that looked like a sheep. However, there was an actual sheep at that time hidden behind the rock.

Does Smith know that there's a sheep in the meadow? As it seems, the answer is negative, in spite of Smith's justified true belief. This is so, because what makes his belief to be true is the actual presence of a sheep in the meadow, but his belief is accidentally true, or *true as a matter of epistemic luck*.

To avoid the factor of epistemic luck the anti-luck epistemologists have tried to solve the Gettier problem by the stipulation that in the terms of the JTB definition the cognitive success of S to believe that P is true is not a matter of luck. This is undoubtedly a plausible stipulation, yet it is not a sound solution to the Gettier problem. As Rodrigo Borges showed for instance (2016: 463), “[the] analysis of luck is of no help to the anti-luck epistemologists for it uses knowledge to explain luck, making this account of knowledge circular.”

As Daniel Whiting puts it, and as it sounds non-optimistically:

We used to think that knowledge is justified true belief. Then Edmund Gettier presented counterexamples to this view which appeared to refute it. Then philosophers spent years, decades even, trying to modify or supplement the view only to see their revised versions face further counterexamples. *Then we gave up trying to say what knowledge is.* (Whiting 2015: 237, my italics)

For now, we see that the notorious Gettier problem – being *central for epistemology*, because it concerns the definition of knowledge – has no commonly accepted solution. Suggested candidates for a solution fail to overcome all the various Gettier cases as counter-examples to the JTB definition. Thus, this naturally transforms the problem into a stumbling block for the epistemologists.

If all that demonstrated so far holds true, it seems that the way to surmount the Gettier problem by a corrective widening of the standard definition of knowledge may not offer a plausible perspective.

### **My Claim**

This is why I will explore another possibility. It is the attempt at demonstrating that Gettier cases are not decisive counter-examples for the JTB definition to be declared as an inadequate one. I'll try to show, in what follows, that there is a problem in accepting the Gettier problem as a genuine epistemological problem. This is so, because even if the Gettier cases intuitively appear to be counter-examples to the JTB definition of knowledge, these intuitive appearances are misleading. Thus, as a non-trivial and unexpected "solution" to the Gettier problem, the statement can be raised that *there is no need for it to be formally solved*.

As far as I am aware, William Lycan was the first author to introduce the unexpected (and astonishing) expression, "the Gettier Problem problem", into the heading of a paper, for the purpose of "explaining what is distinctively wrong with the Gettier project" (Lycan 2006: 150). I'll not follow here, his argumentative reasoning, all the more that he has succeeded in analyzing the peculiarities of different unsuccessful attempts at solving the Gettier problem.

I'll try to provide instead an answer to the question

(Q) "Why do Gettier cases not undermine the JTB definition of knowledge?"

A similar question – "Why are Gettier cases misleading?" – was used as the title of a paper by Moti Mizrahi (2016). The thesis which Mizrahi (2016: 31) defends in his provocative paper is that

[A]s far as Gettier cases are concerned, appearances are deceiving. That is, Gettier cases merely appear to be cases of epistemic failure (i.e., failing to

know that p) but are in fact cases of semantic failure (i.e., failing to refer to x).<sup>4</sup>

I shall try to suggest an answer to (Q) that is different from that of Mizrahi. The gist of it is that *the Gettier cases “are misleading”, because they are based on an implicit self-application of the JTB analysis of knowledge, in order for the conclusion to be always reached that the Gettier-hero (the knowing agent in the Gettier cases) does not know the proposition he believes in, although it proves to be true, and he is justified in believing its truth.*

Even authors who do not hesitate in their support of “the orthodox view in contemporary epistemology” that Gettier cases *refute* the JTB analysis of knowledge, contend that we have a *strong intuition* to say that the respective Gettier-hero (S) does not know that P.<sup>5</sup> As it seems, this pretension for refutation, based on intuition (though a “strong” one), disguises a bit of skepticism. It percolates through the rational doubt that one can hardly reject *a definition* of knowledge by some kind of intuition about its alleged insufficiency.

Let me direct my attention at *the crucial claim* that each Gettier case is aiming at: “S does not know that P”.

My epistemological strategy is finding an answer to the question:

(Q’) How does one *know* that S does not know that P?

It may seem that (Q’) is a strange question. The inventor of a Gettier case contrives the plot of a short story in such a way for the reader to come to the conclusion that S does not know that P. And this looks to be quite convincing at first glance. At a second glance, however, (Q’) must not be grasped and answered in this trivial way. For the intended conclusion in a Gettier case (that S does not know that P) is a “truth” in a fiction tale, or a truth-in-fiction. *If (Q’) is an epistemological question, it has to be supplied with a genuine epistemological answer.* And there is only one way such an

---

<sup>4</sup> See also (Mizrahi 2017).

<sup>5</sup> As Philip Atkins puts it for instance, “I cannot speak for everyone, but I have the strong intuition that Smith fails to know” (Atkins 2016: 381).

answer is to be found: this is the explication of the manner in which the teller of a Gettier case can *know* that S does not know. If the teller does not know this intended conclusion and relies on contrived stories alone, the most she could pretend to tell us is some fabricated scenario.

Furthermore, if the pretention of every Gettier case is that it is an instrument for criticizing a definition of knowledge (being a counter-example to it), then this critical argument must stay on an epistemological level. Otherwise, it gets *outside* of the sphere of knowledge acquisition, and comes to be a mere fabricated story that stays outside of a real cognitive context. *Intuitions* that somebody does not know something are just of this kind. So, it comes out that an instrument having an indirect cognitive nature is used to demonstrate the inadequacy not even of some statement bearing knowledge, but of a definition of knowledge. “The arguments made on the basis of Gettier cases are appeals to intuition, which are themselves a rather controversial sort of argument in philosophy” (Mizrahi 2017: 131).

All this means that *the contriver of a Gettier case must be at the same time involved in a real cognitive situation*. She must be a meta-knowing agent in order to *know* that S does not know that P. Through this reasoning the key epistemological significance of question (Q') is vindicated.

As we have seen from the first Gettier case, its author insists that Smith does not know that (e). I have met the curious objection that Gettier's knowledge that this is so is a second order knowledge, and to this effect Gettier is free to adopt JTB conditions about this second order knowledge, although intending to refute the JTB definition. I agree that Gettier's knowledge is a second order knowledge, since (as I have just pointed out) he is a meta-knowing agent. But second order knowledge is also *knowledge*, as is the first order one; and if we stick to a definition of knowledge, then it must conceptually cover every case of knowledge, be it first order, or second order. Otherwise, the opponents of this claim must clearly show that for having knowledge at an object level, and for having knowledge at a meta-level, one can stick to different definitions of knowledge *per se*. Insofar as such a demonstration is hardly realizable (and not realized for now), one should not use the JTB conditions about

possessing knowledge at a meta-level, in order to refute the JTB definition at an object level.

What remains then is the explication of the way in which a meta-knowing agent does really know that a Gettier hero does not know something, although the hero believes he knows it.

Going back for this purpose to the sheep-in-the-meadow case, let us assume that a friend of Smith who is traveling with him in the same car is playing the role of a meta-knowing agent *who knows* that

(i) Smith does not know that (e'),

where (e') is the proposition "There is a sheep in the meadow".

Smith's friend went along the same road a day ago, and found that there is a rock in the meadow that looks like a sheep. Now traveling together with Smith (and not being the driver), he also managed to see that there is an actual sheep behind the rock several seconds before hearing Smith say that there is a sheep in the meadow. He is thus not astonished to hear Smith's statement that there is a sheep in the meadow, because of the rock's resemblance to a sheep, while knowing at the same time that the statement is factually true.

How then does Smith's friend – in his quality of a meta-knowing agent – *know* the proposition (i)?

His knowledge of (i) is guaranteed by the following three conditions:

(a<sub>1</sub>) (i) is true (since Smith saw a rock, not a real sheep),

(b<sub>1</sub>) He believes that (i) is true, and

(c<sub>1</sub>) He is justified in believing that (i) is true (since he was a direct witness of Smith's misrepresentation of the rock as a sheep).

What comes out is that the way Smith's friend acquires knowledge that (i) is true, is an application of the same JTB procedure that the Gettier case is intending to refute.

By this conclusion, an answer to the question (Q') was provided. One can find a similar answer when analyzing the first original case proposed by Gettier. I have demonstrated that this is so in another paper of mine (Stefanov 2016: 109-110), but there is a need for an additional comment here.

Let me recall that the proposition which Smith does not know in this short story was (e), stating that “the man who will get the job has ten coins in his pocket”. The specification of (Q') now is how Gettier, the inventor of this case, can know that

(ii) Smith does not know that (e), in his quality of a meta-knowing agent, i.e., when involved in a real cognitive situation.

For this purpose, Gettier has to know that Jones has ten coins in his pocket. Suppose that he was a secret eye-witness when Smith was counting them “ten minutes ago”.<sup>6</sup> He is then certain that Smith knows that Jones has ten coins in his pocket. Let us further suppose that he managed to count the coins in Smith's pocket as well (when Smith was buying a cup of coffee at the nearby counter, for instance) and found that Smith also had ten coins in his pocket, and that he had also heard the words of the president of the company, when he was assuring Smith that Jones would in the end be selected for the job for which both men had applied. And when Smith is getting the job, Gettier is correctly coming to know that (ii).

But what does it mean that Gettier *has knowledge* that (ii) in *his* situation of a meta-knowing agent? His knowledge of (ii) is grounded on the following three conditions:

(a<sub>2</sub>) (ii) is true,

(b<sub>2</sub>) Gettier believes that (ii) is true, and

(c<sub>2</sub>) Gettier is justified in believing that (ii) is true.

---

<sup>6</sup> “Smith's evidence... might be that the president of the company assured him that Jones would in the end be selected, and that he, Smith, had counted the coins in Jones's pocket ten minutes ago” (Gettier 1963: 122).



Thus, as a meta-knowing agent, Gettier comes to know that (ii) in the way paved by the JTB analysis of knowledge.

However, a successful application of a cognitive procedure *in order* for its own inadequacy to be demonstrated cannot be used as an argument for reaching this intended aim. This, in its turn, is the answer to (Q), i.e., the answer to the question “Why do Gettier cases not undermine the JTB definition of knowledge?”

This is so, because in spite of the demonstration that the main hero in the Gettier cases does not know something, which he is certain to know, *the knowledge of his not knowing* is controlled by the same JTB analysis of knowledge, the status of which Gettier wishes to undermine.

One could still raise the claim that Gettier cases do formally refute the JTB definition of knowledge. Putting meta-knowledge aside, it suffices to adduce contrived examples of justifications that are misleading. Let me call them “wrong justifications”, although this phrase is somehow awkward. The possibility for wrong justifications is put forth as an argument against the adequacy of the JTB definition. Well, the whole history of human knowledge, and of what we call rational knowledge too, is full of wrong justifications. These contingent cases, however, in no way cancel *the very need for justification*. Justification remains as a necessary conceptual requirement for having knowledge. Accidental cases of wrong justifications cannot delete the *conceptual* requirement for knowledge to be defined as justified true belief. *Wrongly* justified true belief may simply not be accepted as genuine knowledge.

Accidental cognitive traps are merely influences of misleading situational factors that could be *principally avoided*. They do not entail a *conceptual* failure of the JTB definition of knowledge *per se*. Their emergence is contingent and has nothing to do with the conceptual setting of the JTB definition of knowledge. This definition – as bearing *an understanding* of what is it for S to know that P – must not be blamed for putative knowledge failures.

There certainly are different kinds of justifications, which are valid for different types of cognitive situations. And the more strictly the pertinent

requirements for a justification are followed, the less is the room left for accidental cognitive traps leading to wrong justifications.

What follows in the end is that the JTB analysis of knowledge *is not refuted at a conceptual level* by the fabrication of Gettier cases. The presentation of the latter as counter-examples to the JTB definition of knowledge is backed up by this same definition.

Hence the Gettier problem ought not to be looked upon as a stumbling block for epistemologists.

## CHAPTER 2

# DO SELF-REFERENTIAL INCOHERENT THEORIES REFUTE THEMSELVES?

### **The Traditional Positive Answer**

A theory (or a statement) is said to be *self-referential* if theories (statements) and their basic features are included in its subject matter, or to put it in other words, if it is about theories (statements) and is included in its own domain. Thus, for example, the statement that “Every statement is false” is self-referential, because it refers to all statements and it is itself a statement.

A self-referential theory may, or may not conform to its own criteria of validity or acceptability. The statement that “Every statement is false” certainly does not. Insofar as it refers to all statements, it also refers to itself. Let us assume that this statement is true. Then it follows, because of what it says, that it is not true. It thus runs counter to the assumption of its own truthfulness. A self-referential theory which does not satisfy its own criteria of validity, or breaks up in some way its own requirements, is said to be *self-referentially incoherent (inconsistent)*.

Self-referential incoherence is traditionally considered in its role of a powerful method of refutation.

Sometimes, on learning that I am a philosopher, a non-philosopher who has gained the impression that it is difficult to get agreement in philosophy will tease me with the idea that philosophy can never definitely establish anything. I then like to point out to my would-be tormentor that he has just unwittingly provided me with all the material I need to refute that very thesis. For the idea that it is impossible definitely to establish any philosophical thesis is itself a substantive philosophical thesis. Therefore we *can* definitely establish that if my interlocutor ever definitely established his thesis that it is impossible definitely to establish any philosophical

thesis, then in establishing his thesis, he would have refuted it. Thus we can definitely establish that one can never definitely establish that one can never definitely establish anything in philosophy. So we can definitely establish something philosophical. (Zuboff 2015).

The just quoted excerpt is a nice consolation for the philosophers. And I surely would not make up my mind to oppose its clever setting.

But what can further attract philosophical interest is the analysis of the pretention of a thesis that gives a clear positive answer to the title's question:

(R) *Do self-referential incoherent theories refute themselves?*

It is an established, though non-obtrusive, tradition (R) to be accepted as a valid thesis without any relevant critical comments. This means further that as a normative claim, and if ascertained for some self-referentially incoherent theories, (R) could be used as a *trustworthy method for the refutation of all such theories*.

Thus, for instance, the famous Carl R. Kordig (1983) has avowed the universality of this method and declares, by applying it, some well-known epistemological views (theories about theories) to be false. According to him, Toulmin's evolutionary epistemology must be rejected, because it appears to be a false theory about theories just on the grounds of (R).<sup>7</sup>

A. Zuboff (2015) goes along with the same line of reasoning, and declares the refutability of Wittgenstein's description of language as a game, as well as of the position of deconstructors of texts with respect to their own texts.

---

<sup>7</sup> Other epistemological views that Kordig (1983: 208) finds to be self-referentially inconsistent, and thus self-refuting, are the Quine-Duhem thesis, Salmon's account of factual content, Quine's ontological relativity, the claims that not everything has an explanation and cause, and some others. In fact, Kordig puts forward a bilateral claim which runs to the following: If a self-referential theory is incoherent, then it is necessarily false; its denial, however, is self-validating and hence necessarily true. I have hopefully shown that this claim may face some specific difficulties (Stefanov 1998: 135-136).

*The aim of this chapter* is to defend the claim that the method of refutation, based on (R), is not always reliable, and that at least some seemingly self-referential incoherent theories ought not to be looked upon as self-refuting.

I don't maintain any doubts that *from a formal logical point of view*, and when all theories are taken to be expressed within a common language, (R) is undoubtedly a valid thesis. Let us consider for example the statement:

(S) Every statement expressed in English is false.

(S) refers to all statements expressed in English, and it is itself a statement formulated in the same language. If all such statements are false, then as one of them, (S) must also be necessarily false. But if so, its negation comes to be true, which means that at least one sentence in English is true. The latter conclusion shows that (S) is incoherent, what it states is not true, and thus, according to (R), it is self-refuting.

Let us now have a look at another example of a statement that allegedly refutes itself:

The renowned pragmatist W. V. Quine famously claimed that in our proper theories of the world (...) *no statement is unrevisable*. But isn't this statement, Quine's statement that *no statement is unrevisable*, meant by Quine to *be* unrevisable, that is, to represent always what is true about revisability whether or not there were pragmatic grounds for maintaining that supposed truth? ... Quine's statement refutes itself... [I]f next we try to think of his statement as *unrevisable*, we find that it would of course at the same time be false since it says there are *no* unrevisable statements. (Zuboff 2015).

But does Quine's statement really refute itself?

I am inclined to say "Yes, it refutes itself", if this statement is construed as one formulated within a definite language, all of the sentences of which, expressing different statements, refer to one common domain. Thesis (R) is a good method for refutation in this case, and can be safely applied in a formal logical reasoning. Indeed, if no statement formulated within a

language is unrevisable, then the very statement bearing this message, as being self-referential, falls under this factual situation. This leads to the acceptance of its revisability, and hence to its falsity, since the statement must be either true, or false.

However, I can also say “No, Quine's statement does not refute itself”, because it appears to be not properly self-referential. By saying that “it is not properly self-referential”, I mean that this statement is formulated in an epistemological context, which differs from the one encompassing the set of all statements to which it refers. Thus, its prescription directed to the latter statements may not (directly) refer to itself.

Indeed, in the case being considered, all of the referent statements belong to “theories of the world”, that is to say, to factual scientific theories. However, Quine’s statement is a theory about scientific theories, and not a statement belonging to their sphere of validity; it is not a statement describing any fragment of, or state of affairs in the natural world.

If so, Quine’s statement does not belong to the set of statements about which it prescribes something. It is not properly self-referential. Its prescription might prove to be valid for the statements it refers to, being at the same time invalid for Quine’s statement itself.

Hence, (R) cannot be used as a method for the refutation of Quine’s statement.

The conclusion just reached is a counter-example to thesis (R) taken in its formal logical pretention. It is enough for the defense of my suggested claim that the method for refutation, based on (R), is not always reliable.<sup>8</sup>

---

<sup>8</sup> Looking back at Quine’s statement, this same conclusion does not, however, mean that the statement is necessarily true, and on these grounds, it is unrevisable. But *even if* this statement is really revisable, *its putative revisability is not due to the application of (R) to its semantical content*. Its revisability may be contended on epistemological grounds, exploiting arguments eventually provided by some kind of strong metaphysical realism.

## Mavrodes' Broad Answer

George Mavrodes does not examine *why* a given self-referential theory is incoherent, but the problem what *if* it is such a theory.

We shall want to ask whether various theories which do *not* satisfy themselves are vulnerable in different degrees to the force of this kind of refutation. (Mavrodes 1985: 66)

Beyond any doubt, posing this question implies some kind of amended version of the thesis (R), or some other solution to the title question of this chapter.<sup>9</sup> In fact, Mavrodes does not supply such a worked-out solution. Nevertheless, he shows that the problem itself (about the refutability of self-referentially incoherent theories) is not a pseudo-problem, and that its solution depends on some special presuppositions which he calls "principles of evaluation". I shall make use of his two straightforward examples here (Mavrodes 1985: 66).

Let us examine the proposal:

(T) Truth is beauty.

Mavrodes supposes that in criticizing (T) we have managed in some way to show that (T) does not satisfy itself. This would mean that (T) is not beautiful. From here we can infer that if (T) is true, then (T) is not true. But the last conclusion could present enough grounds for the refutation of (T) only in combination with the following principle of evaluation:

(P<sub>1</sub>) For the analysis of truth, it is a desideratum that the analysis be true.

And insofar as (P<sub>1</sub>) seems to be quite arguably a principle, (T) must be rejected as an unsatisfactory analysis of truth.

If we apply the same argumentation, however, in order to ascertain the refutability of another, though similar claim, the result will be different. The claim is;

(F) Falsehood is ugliness.

---

<sup>9</sup> Kordig's paper (1983) is not among the references of Mavrodes' paper (1985).

Let us assume that we can show the self-referential incoherence of (F) as well. By using exactly the same way of reasoning as in the preceding example we can reach the conclusion that (F) is not false. To refute (F) will suffice to introduce the evaluation principle

(P<sub>2</sub>) For the analysis of falsehood, it is a desideratum that the analysis be false.

But insofar as (P<sub>2</sub>) is not at all an acceptable principle, the self-referential incoherence of (F) itself cannot present sufficient grounds for the refutation of (F).

By the way, this second example suffices to show that we cannot rely on a general principle of evaluation of the form:

(P<sub>3</sub>) For any property  $\phi$ , it is a desideratum that the analysis of  $\phi$  should itself have the property  $\phi$ .

That is why we must look for a convenient and acceptable principle of evaluation, each time deciding whether a self-referentially incoherent theory should be rejected as unsatisfactory. This requirement bars Mavrodes from getting at a more detailed solution to the problem “What if a theory is self-referential and incoherent?”, and makes him content with the following general conclusion:

In some cases, most notably those of truth and falsehood, the subsequent argument is straightforward and powerful. But in other cases the argument soon leads us into complexities which leave the force of the intended refutation much more in doubt. (Mavrodes 1985: 72)

For the sake of precision, someone could raise the objection that claims (T) and (F) are not convenient examples for the analysis of self-referential incoherence. And she will certainly be right, insofar as both of these claims are semantically incorrect. Truth and falsehood are not the qualities which are supposed to fall into the range of applicability of the corresponding predicates “beautiful” and “ugly”. But even if this is so, this circumstance does not run counter to Mavrodes’ warning that *the thesis (R) alone is not a sufficient condition for the refutation of every self-referential incoherent theory*. A sufficient condition for this is the



combination of (R) with a relevant principle of evaluation. Mavrodes' kind of solution to the title question of this chapter thus stays in harmony with my claim that "the method of refutation, based on (R), is not always reliable, and that at least some seemingly self-referential incoherent theories ought not to be looked upon as self-refuting".

Mavrodes' kind of solution, however, is not an explicit one, since it relies on the fabrication of principles of evaluation, having different degrees of relevance for each separate case of a self-referential incoherent theory.

### **My suggested solution**

I suggest another kind of solution, which is based on an attempt at *providing an explicit typology* of self-referential incoherent theories, which are susceptible to refutation on the basis of (R).

I shall denote by (A) the class of self-referential, incoherent, *singular*, and *semantically correct statements*. An (A)-type statement is represented for instance by the sentence:

(1) This very sentence is in Spanish.

All three requirements are satisfied by (1). It presents a singular statement, and the demonstrative reference of its subject is (1) itself. Furthermore, (1) is semantically correct, because the reference of its subject *is* included in the range of applicability of its predicate. Indeed, the predicate "Spanish" can well be referred to sentences, and (1) is surely a sentence. But the obvious inconsistency demonstrated by the presented statement makes it false and thus it is self-refuting.

(B)-type incoherent theses are *singular* statements resembling those from class (A), but *semantically incorrect*. It is namely this semantical incorrectness that is responsible for their incoherence. A (B)-type sentence is:

(2) This very sentence is green.

Linguistical structures – and (2), as a sentence, is one of them – do not naturally fall into the range of applicability of the predicate "green". That

is why (2) is semantically incorrect. And since it is incoherent and self-referential, it is self-refuting.

Unlike (A) and (B), class (C) covers all *general* statements and theories which are *properly included in their own domains*, and fail to satisfy themselves. (C)-type theories are clear instances of self-referential incoherence, and are refuted merely because of this. Such is e.g., the general statement:

(3) Every statement is false.

Such is also the Quine-Duhem epistemic thesis taken in the form:

(4) No hypothesis whatsoever can be conclusively falsified.

The negation of (4) is itself a hypothesis. That is why it too cannot be conclusively falsified. The negation states that some hypotheses can be conclusively falsified.

Thus, for all we know, some hypotheses can be conclusively falsified... It is therefore not true that we know that no hypothesis whatsoever can be conclusively falsified. Thus, if we know that no hypothesis whatsoever can be conclusively falsified, it follows that we do not know that no hypothesis whatsoever can be conclusively falsified... Quine-Duhem epistemic thesis is self-referentially inconsistent and hence ought to be rejected. (Kordig 1983: 209)

There is a very special class of statements, which I shall denote by (C'). All statements within (C') are *of the type of the Liar Paradox*:

(5) This very sentence is false.

(6) Everything said by me is false.

These are examples of self-referential incoherent claims. Nevertheless, I do not unite (C) and (C') because of the presence of singular statements like (5) in (C') and their absence in (C). And I do not unite (C') and the class (B) for another reason, in spite of the fact that (5) and (6) cannot pass the standard procedure for semantical correctness (Martin 1976: 298-

303).<sup>10</sup> This is because the (C')-statements refer to another convention for semantical incorrectness: namely that each sentence which explicitly mentions itself (appearing at the object- and meta-language level simultaneously) must be looked upon as semantically incorrect. However, the problem is not in accepting the convention, but in the fact that it is too discriminative. Together with all incoherent sentences that explicitly mention themselves, it rules out all self-referential *coherent* sentences as well. Such is for example the declaration that "This sentence is in English". That is why the convention ought to be specified for the purpose being here pursued. I do not see any other possibility for an appropriate specification of the set of semantically incorrect statements of the type considered, except the following proposal: we must have in mind all sentences which explicitly mention themselves, and contain as a predicate the self-qualification "false". Thus, the separation of the class (C') obtains its justification.

Now I come to the last class of allegedly self-referential incoherent theories which, in contrast to (A), (B), (C), and (C'), *are not susceptible to the refutational strength of (R)*. The biggest problem lies namely within this class of theories which I shall denote as (D)-type theories for short. There are *theories about the nature and functions of scientific theories* which exhibit incoherence, as in the already mentioned case (at the beginning of this chapter) with Toulmin's evolutionary theory. Their contingent refutations are of great interest to philosophers, insofar as *their metatheoretic character has an epistemological lineament*.

The alleged incoherence within Toulmin's theory, according to Kordig, is the discrepancy between its evolutionary normativeness, and the necessity for it, if it should itself conform to the latter, to evolve into a non-evolutionary theory.

Let me recall the examples used by Toulmin in his *Human Understanding*, to illustrate his idea about the change in the general intellectual goals and explanatory ideals in the course of scientific growth.

---

<sup>10</sup> Instead of (6) Martin considers the statement "Everything said by a Cretan is false (said by a Cretan)". There is no need, indeed, to show the inconsistency equivalence of Martin's example and (6).

The first one concerns Max Planck's analysis of the altering explanatory demands which have historically guided the development of physics, presented in the exchange of papers between him and Ernst Mach, printed in the *Physikalische Zeitschrift* for 1910-11 (Toulmin 1977: 232). Toulmin points out, as another example, the debate among representatives of classical genetics (Avery and his colleagues) and physicists/biologists (Delbrück and others) in the years between 1944 and 1953. The third example is taken from quantum physics, and concerns J. Schwinger's and G. Chew's approaches to the explanation of the nature and character of the elementary particles.

Here again, the subject faces theoretical difficulties that call, not for more elegant mathematics or more ingenious experiments alone, but rather for a strategic reappraisal of basic aims and explanatory ideals. (Toulmin 1977: 234-235)

These examples from the history of science clearly outline *the pretension of Toulmin's epistemological view*. This refers to *scientific theories* and how new theories supersede old ones because of changing intellectual goals and ideals. The intellectual goals in science and the general ideals of scientific understanding are the factors that historically evolve, *but not* the epistemological goals that guide Toulmin's own evolutionary theory. In this sense, I would say, *Toulmin's evolutionary theory is not properly included in its own domain*. It is about scientific theories, and not about epistemological views. Hence, in contrast to (C)-type self-referential incoherent theories, which are properly included in their own domains, Toulmin's evolutionary epistemology cannot be refuted on these same grounds.

Of course, an immediate objection may be raised that even if Toulmin's pretensions do not spread over epistemological metatheories, this is, and could be no argument to change the situation. Because theories are human cognitive constructions, and *being such*, they must exhibit a common nature and common basic features independently of whether they are theories about regions of the outer world, or theories about theories. To this effect, Toulmin's metatheory must be necessarily considered as

properly included in its own domain, and hence, should be rejected as self-referentially incoherent.

Is this objection correct? It is, except with respect to one crucial assumption; the general assumption that a metatheory about theories, because it *is* a theory itself, must necessarily possess the conceptual features of the theories it describes. A special instance of this general assumption (reporting on a single property) is the evaluation principle ( $P_3$ ) pointed to by Mavrodes. But if even ( $P_3$ ) is not an acceptable principle, what about the general assumption which functions rather as a prejudice, than as an established normative rule?

Let us examine another, far simpler, but no less paradigmatic example of a (D)-type theory:

(7) All theories have a restricted sphere of validity.

And let us assume that (7) tells us the truth. Then, insofar as it is itself a theory about theories, it follows that its own sphere of validity is restricted. But since it is about theories, this implies that some theories are *not* characterized by a restricted sphere of validity. So, (7) runs counter to its own assumption. It is self-referentially incoherent.

Can (7) be refuted because of this? Yes, if we involve an additional premise in the form of ( $P_3$ ) which states that for an adequate analysis of the sphere of validity of theories the analysis itself should possess the same type of validity. But the cognitive necessity of such a premise has never been proved. Thus, it comes out that the alleged self-referentially incoherent theory about theories (7) is not refutable on the grounds of the thesis (R).

It becomes clear in the end that (R) has a real refutational strength for self-referential incoherent theories covered by the classes (A), (B), (C), and (C'), but cannot be directly applied to (D)-type theories. This in no way means that the theories of this type are not liable to criticism. However, the criticism is based on various arguments, and stays out of the grip of the thesis (R).

## CHAPTER 3

### WHAT IS NON-CLASSICAL THEORY?

The aim of this chapter is for an arguable answer to the title question to be reached.

If I succeed in so doing, it might be expected that the answer ought to provide a necessary and sufficient condition according to which a given scientific theory could be accepted as a non-classical one.

I concede at the outset that I cannot suggest such a complete answer, since I cannot provide its first “necessary” component. I shall try, however, to suggest *a sufficient condition for a scientific theory to be looked upon as a non-classical one*.

But why does the title question deserve philosophical attention? The answer (at least) to this second question is not difficult. It is a fact that the quality of “non-classicalness”, attributed to scientific theories, has a historical usage.

R. Descartes laid the beginnings of a new philosophy which buttressed the new science of his time. It could be qualified as non-classical, regarding the previous system of thought. This notwithstanding, Descartes is well-known nowadays as a prominent representative of *classical* rationalism. Non-Cartesian epistemology was taken by some thinkers to be a non-classical trend in philosophy.<sup>11</sup>

Quantum mechanics was proclaimed to be a genuine non-classical theory about the nature of the micro-world, but since the 1990s it has often been referred to as “classical quantum theory”, in order to be differentiated from quantum field theory, superstring theory, or later quantum achievements

---

<sup>11</sup> About non-Cartesian epistemology see, for example, Bachelard (1934: ch. 6). Till 1984 the book underwent 16 editions in France.