A Priori truth in the natural world : a non-referentialist response to Benacerraf's dilemma
Novák, Z.

Citation
Novák, Z. (2010, June 23). *A Priori truth in the natural world : a non-referentialist response to Benacerraf's dilemma*. Retrieved from https://hdl.handle.net/1887/15729

Referentialist Responses to Benacerraf's Dilemma I: Deflationist Construals of Truth

Introduction

Deflationist theories of truth can be characterised by their allegiance to the claim that truth is not a substantive property whose nature and relation to our epistemic capacities could be specified by a proper theory of the subject. In this, deflationism can be contrasted with all realist and anti-realist theories of truth. Alternative versions of deflationism endorse various further claims about the cognitive function of our concept of truth.[1] Most deflationists today believe that the concept is merely a logical device, which does not stand for a property whose nature could be further specified. When we apply the predicate 'is true' to certain thoughts or their linguistic expressions, we do not attribute a real property to these representations, but instead we assert something that could as well have been asserted by endorsing the relevant thoughts or sentences.

By adopting a deflationist theory of truth, an advocate of standard referentialism can provide a *prima facie* acceptable response to Benacerraf's epistemological challenge in the

---

[1] There are several classifications of deflationist accounts in the literature. The main versions that are usually distinguished include Ramsey's Redundancy Theory (Ramsey (1927)), Tarski's Semantic Theory (Tarski (1944)), Strawson's Performative Theory (Strawson (1950)), Quine's Disquotationist Theory (Quine (1970)), Grover's, Camp's and Belnap's Prosentential Theory (Grover, Camp and Belnap (1975)), and Horwich's Minimalist Theory (Horwich (1998b), Horwich (2005)). Further proponents of deflationism include Ayer (1935), Wittgenstein (1953), Leeds (1978), Fine (1984), Soames (1984), Soames (1997), Field (1986), Field (1994), Williams (1986), Loar (1987), Brandom (1988), Brandom (1994).

semantics of our paradigm *a priori* discourses. As it has been shown in chapter 3, in its most generic form, the epistemological argument questions the adequacy of a referentialist *and* realist construal of truth in the semantics of discourses in which we are supposed to acquire knowledge or reliable beliefs about causally inert domains. By abandoning the realist tenet, a deflationist can subscribe to the standard referentialist construal of truth without falling prey to Benacerraf's argument. Clearly, if truth is not a substantive property, then referentialism about truth does not imply the (thick) existence of the subject matter of true beliefs either. In a deflationist semantical framework, referentialism can be maintained independently of what can be truly said in metaphysics and epistemology. In the semantics of our paradigm *a priori* discourses, this means that our beliefs can be regarded as true or false in the standard referentialist sense independently of what can be said about the metaphysical status of the relevant abstract subject matters, and the ways in which we acquire knowledge or develop reliable beliefs about them. A deflationist may admit that she has no suitable explanation of how we acquire knowledge of causally inert domains. Nevertheless, in view of the alleged autonomy of semantics from metaphysics and epistemology, this failure is not supposed to interfere with her belief that our notion of truth is best specified in the standard referentialist way.

In this chapter, I shall examine the viability of this deflationist version of standard referentialism, and argue that the indicated advantages of the theory are cancelled out by the fact that a deflationist can provide no suitable explanation of the objectivity of truth. This is because in absence of a real distinction between the situation in which truth does characterise a particular representation and that other in which it does not there seems to be no way to understand how the obtaining of either of these options could be conceptually independent of what anyone ever believes about this issue. In other terms, a deflationist conception of truth does not tell us why none of us is ever conceptually prevented from committing an epistemic mistake concerning the applicability of representations. A realist

construal of truth provides a natural explanation of this phenomenon: the conceptual possibility of these epistemic mistakes is a consequence of the reality of truth. If so, then by applying the predicate 'is true' we attribute a real property to our representations. Moreover, this property seems to have an underlying relational nature as well, which can be characterised as a sort of correspondence. According to this construal, truth is the property of having suitable semantic relations to some conditions that obtain in the actual world, or having a semantic content that corresponds to some fact in the actual world.

Since the arguments that I shall advance attack a general feature of deflationism, there is no need to discuss the various formulations of this programme. Instead, in section 1, I shall focus on the best elaborated deflationist conception in the current literature, Paul Horwich's minimalist account of the concept of truth. After a brief presentation of his position, I shall argue that although many of his core insights about our concept of truth seems to be well-considered and adequate, nevertheless his negative tenet that by applying this concept we do not attribute a substantive property whose nature and relation to our epistemic capacities can be specified by a proper theory of the subject is clearly incompatible with one of the key distinctions of any plausible theory of meaning and truth (including his own "*use-regularity* conception of meaning"). The distinction in question is that between the actual and the correct use, or between our belief in the correctness and the objective correctness of the use, of various sorts of representations. The incompatibility obtains, because the distinction presupposes what Horwich's minimalism, together with other forms of deflationism, denies: namely, that truth is a real property, which characterises our truth-apt representations independently of whether anyone ever knows of the obtaining of these particular states of affairs.

Having developed my argument against the deflationist conception of truth, in section 2, I shall defend the received realist (i.e. correspondence) theory of this semantic property against the so-called "slingshot argument", which is supposed to

demonstrate that all true beliefs refer to the same thing, and therefore cannot be made true by the obtaining of their own *separate* truth conditions. Although the argument's conclusion is compatible with a realist construal of truth, and also with the identification of truth with the property of having suitable semantic relations to some (namely: all) conditions that obtain in the actual world, nevertheless it undermines the received realist assumption that our true beliefs may refer to, and be made true by, different facts in the world. After presenting the argument in its simplest and most transparent form, I shall show that two of its background assumptions rely on a highly coarse-grained (Fregean) notion of reference, which is different from the one that we use when specifying the semantic content of our representations along the standard referentialist lines. If we replace the former collapse-generating concept with the latter fine-grained notion in these assumptions, then the argument breaks down, and no longer challenges the adequacy of the received realist (i.e. correspondence) theory of truth.

1. Minimalism and the Explanation of the Objectivity of Truth

Paul Horwich's minimal theory of truth is the most fully elaborated and widely discussed form of deflationism today. In the second edition of his book on the concept of truth, he introduces his account as a reaction to two misconceptions entrenched in present-day philosophy: "first, that truth *has* some hidden structure awaiting our discovery; and secondly, that hinging on this discovery is our ability to explain central philosophical principles […], and thereby to solve a host of problems in logic, semantics, and epistemology".[2] The principles whose explanation seems to call, according to many philosophers, for a substantive notion of truth include, for instance, that truth

is the aim of science, that true beliefs tend to facilitate the achievement of practical goals, that truth is preserved in valid reasoning, or that to understand a sentence is to know which circumstances would make it true. According to Horwich, the presented misconceptions are rooted in our belief that the predicate 'is true', like other familiar predicates, such as 'is magnetic' or 'is diabetic', also designates a certain complex feature of the world whose underlying structure could be revealed by a succinct philosophical or scientific analysis.

In contrast to this belief, Horwich maintains that "the truth predicate exists solely for the sake of a certain logical need".[3] Occasionally, we may want to endorse propositions that we cannot directly identify by using their name, either because we are ignorant of what those propositions are, or because they are so many that we cannot grasp them all in our mind. In such cases, we can express our attitude by applying the truth predicate to the available descriptions of the relevant propositions (e.g. "the proposition that Peter argued for yesterday is true", "every proposition of the form *p or not p* is true"). Beyond this logical function, there is no role for our concept of truth to play, and a proper account of this concept should say no more than what is necessary for explaining this role. It is not that the concept *cannot* figure in the explanations of the earlier cited principles. It can also appear in accounts of such truth-related phenomena as knowledge or the nature of various types of fact. Horwich's point is rather that the theoretical role that our concept of truth plays in the latter sorts of explanation can be derived from the former logical function, and therefore a theory which can account for that function must be able to carry out these further explanatory duties as well.

The minimalist conception advocated by Horwich is that our concept of truth can be fully characterised by the statement that "the law governing its use is that we are prepared to provisionally accept any instance of the schema, '<p> is true ↔

---

[2] Horwich (1998b), 2.

[3] Horwich (1998b), 2.

p'".[4] Accordingly, a theory of truth should contain nothing more than the uncontroversial instances of the equivalences schema 'It is true *that p* if and only if p'.[5] With this minimal theory in mind, his main purpose is to demonstrate that virtually all phenomena whose explanation was traditionally meant to call for a substantive account of truth can be explained on the basis of the above minimal assumption alone.[6]

How does this minimalist project relate to the naturalistic construal that I wish to advocate in this work? On a short reflection, it may become clear that they are plainly incompatible with each other. First, while on the naturalistic construal to be proposed, truth and falsity are substantive real properties of thoughts and their linguistic expressions that characterise these bearers in virtue of what the latter mean what actually obtains in the spatiotemporal world, the minimalist conception denies that our concepts of truth and falsity stand for such properties in the world. Second, while the naturalistic construal is motivated by the conviction that its substantive assumptions about truth are necessary for explaining all observable phenomena involving this semantic property, the minimalist conception is based on the belief that the explanatory duties in question can be carried out even if we reject those assumptions.

In what follows, I shall argue that the minimalist tenets mentioned in the previous paragraph are equally false. First, I shall show that if we endorse Horwich's positive claim about the notion of truth, namely that it can be truly characterised by the instances of the equivalence schema, and we accept his invitation to adopt a "use-regularity" conception of meaning, which

nevertheless preserves the conceptual difference between the actual and the correct declarative use of various sorts of representation, then we can derive the realist idea that truth is the property of having suitable semantic relations to conditions that obtain in the actual world. What this demonstration will show us is that Horwich's negative, deflationist claim about the notion of truth cannot be reconciled with the two positive commitments mentioned above.

As we have seen, according to the minimalist conception, a theory of truth should contain nothing more than the uncontroversial instances of the equivalence schema 'It is true *that p* if and only if p'. As Horwich also recognises, in order to be informative, the account must presuppose that we can understand the right-hand side of these instances without relying on the notion of truth. In other terms, the theory can provide an informative characterisation of truth only if our understanding of propositional contents is not truth-conditional.[7] The way in which Horwich ensures the informativeness of his minimalist conception of truth is by replacing the truth-conditional with a use-theoretic account of meaning.[8] According to this account, the meaning of a term derives from its use, and the acquisition of meaning requires merely the recognition of some basic regularities in use, which can be achieved without reliance on the concept of truth.

Of course, meaning cannot simply be meant to reduce to actual use, since this would, among other things, exclude the conceptual possibility of incorrect use. A further fact about meaning to be taken into consideration is that, once it has been established, no one is conceptually prevented from committing an epistemic mistake while deliberating over the actual applicability of a certain representation, since no one's judgement about this issue is constitutive of the *objective* applicability of that representation. Accordingly, a proper theory of meaning must

---

[4] Horwich (2005), 26.

[5] Horwich (1998b), 7, 19-20.

[6] As Horwich observes, in the case of those *explananda* which concern a relation between truth and an other thing, such as knowledge acquisition, "it is perfectly proper to make use of theories about these other matters, and not to expect all the explanatory work to be done by the theory of truth in isolation. Horwich (1998b), 7. So, on his view, the virtue of minimalism is "that *it provides a theory of truth that is a theory of nothing else, but which is sufficient, in combination with theories of other phenomena, to explain all the facts about truth*". Horwich (1998b), 24-25.

[7] Horwich (1998b), 68-71.

[8] For a detailed presentation and defence of this theory, see Horwich (1998a) and Horwich (2005).

construe this semantic property in a way that accounts for the conceptual possibility of incorrect application in the case of each epistemic agent. If the factual ground of this possibility is, indeed, that the correctness of a certain use is independent of our actual beliefs about this issue, then the position at which the previous constraint leaves us is that a proper theory of meaning must construe the conditions of correct use in a realist way.

On Horwich's view, the meaning of a term reduces to a (non-semantic) use or acceptance property which characterises the term due to those *basic regularities* that underlie its application. Accordingly, two terms are supposed to have the same meaning if and only if they instantiate the same use property (i.e. their use displays the same basic regularities). Now, if this "use-regularity" account is supposed to meet the above adequacy condition for a theory of meaning, then it must affirm that the conditions under which the use of a certain representation displays a certain basic regularity obtain independently of our actual beliefs about this issue (i.e. they are to be construed in a realist way).[9]

But how do we arrive from this realist constraint upon Horwich's theory of meaning at a realist challenge to his minimal theory of truth? To see this, we have to return to his conception of how the two theories relate to each other. As we have seen, his view is that our concept of truth can be fully characterised by the statement that we are prepared to apply the predicate 'is true' in the context of the proposition *that p* if and only if *p*. Now, supposing that this conception is correct, and also that the

conditions under which the use of 'is true' in the context of the proposition *that p* can be regarded as objectively regular must be interpreted in a realist manner, we can derive that the right-hand side of the instances of the equivalence schema must be understood also in a realist way (i.e. specifying conditions whose obtaining or absence is independent of what anyone ever believes about this circumstance).[10] According to this reading, what Horwich's minimal theory specifies (again, not necessarily by referring to them) are the objective declarative use conditions of our predicate 'is true' in the context of any conceivable proposition.[11] In possession of this theory, we will hold, indeed, the same capacity as in possession of our concept of truth: namely, for each *understood* proposition, we will be able to entertain an idea of those (non-epistemic) conditions whose obtaining in the world is necessary and sufficient for the correct declarative applicability of our predicate 'is true' in the context of that proposition.

The first important claim that I wish to derive from the previous conclusion is that even if we suppose, I think correctly, that the acquisition of meaning does not require the possession of the concept of truth (for instance, because we can recognise and follow the basic regularities characterising a sentence's use

---

[9] In *Truth* Horwich addresses the charge that a deflationist perspective leads inevitably to relativism: to the idea that there is no such thing as objective correctness. Horwich (1998b), 52-53. His response, in my view correct, to this challenge is that the denial of a substantive realist account of truth should not be confused with the endorsement of a substantive anti-realist construal of this property, which may, indeed, give rise to "an extreme form of relativism in which it is supposed that truth is 'radically perspectival' or 'contextual' or something of the sort". Horwich (1998b), 53. The current challenge, however, is not that a deflationist semantics denies the existence of objective correctness and entails radical relativism, but instead that it cannot account for the objectivity of correct use.

[10] As it was emphasised earlier in this work, a realist construal of the declarative use conditions of certain representations need not commit us to the realist construal of what these representations purport to be about. This is because truth need not be understood in referentialist terms. Consequently, the current claim that the right-hand side of the instances of the equivalence schema have to be read in a realist way has no implication of the metaphysical status of the subject matter of these propositions either. Of course, if one adopts a referentialist construal of truth, then the commitments in question will follow.

[11] The clause "not necessarily by referring to them" is meant to emphasise the idea that the equivalences stated by the instances of the equivalence schema hold only because the right-hand sides of these instances need not be read in a referentialist way. In the case of propositions about abstract domains, for instance, one can endorse, as I would certainly wish to do, that 'the proposition *that two plus two equals four* is true if and only if two plus two equals four' without thereby subscribing to the referentialist claim that the truth conditions of this arithmetical proposition are to be specified in terms of the abstract subject matter of the right-hand side of this instance of the equivalence schema.

without understanding that the conditions whose obtaining or absence informs these regularities may also be regarded as the conditions of the sentence's truth), since the conditions specified by the right-hand side of the instances of the equivalence schema are the correct declarative application conditions of our predicate 'is true' in the context of the relevant propositions, they can be equally understood as the truth conditions of these truth-apt representations.[12] Having said that, with the former claim that these conditions should obtain independently of our actual beliefs about this issue, we may conclude that, contrary to the deflationist tenet, the truth conditions of our truth-apt representations are to be interpreted in a substantive realist way (i.e. that truth must be regarded as a real property after all). Moreover, if the truth conditions of our thoughts are indeed identical with the correct declarative use conditions of these representations, then truth must be identical with the property of having suitable semantic relations to conditions that obtain in the actual world.[13]

Second, we may notice that the derivation of the above realist construal of truth was based on three major premises. First, we granted that our notion of truth can be truly characterised by the instances of the equivalence schema. Second, we accepted that the meaning of a term derives from its use, and

---

[12] It may be worth noting that Horwich does not query the correctness of this observation. What he denies is merely that understanding derives from knowledge of these conditions *as* conditions of truth. Horwich (1998b), 69-70. An early recognition of this consequence of the acceptance of the disquotation or equivalence schema can be found in McDowell (1981): "There is a truistic connection between the notion of the content of an assertion and a familiar notion of truth … the connection guarantees, as the merest platitude, that a correct specification of what can be asserted, by the assertoric utterance of a sentence, cannot but be a specification of a condition under which the sentence is true." McDowell (1981), 229. Further endorsements of this point can be found in Wright (1993), 18-19, and Miller (2002), 364-366.

[13] Since, on this realist construal, the truth conditions of our beliefs and utterances are supposed to obtain in reality independently of the occurrence of these representations, truth may also be characterised as a substantive relation ("correspondence") between the latter bearers and the former aspects of reality.

the acquisition of meaning requires merely the recognition of some basic regularities in use, which can be achieved without reliance on the concept of truth. Finally, we maintained that, in order to account for the objectivity of correct use, a proper theory of meaning must construe the conditions of regular applications in a realist way. Since Horwich's minimalist conception contradicts the realist conclusions presented above, at least one of the former three premises must be incompatible with his position. Given that the first two assumptions are explicitly endorsed by him, what he presumably rejects is that a theory of meaning can be adequate only if the conditions of correct use are to be understood in a substantive realist way. In this, however, he could be right only if either a proper theory of meaning did not have to account for the objectivity of regular use, or it could do this without the adoption of the suggested realist tenet. Since, on my view, neither of these conditions obtains, I conclude that Horwich's presumable rejection of the above realist tenet is a mistake, which undermines the explanatory adequacy of his theory of meaning. Moreover, since his minimal theory of truth implies that truth can be understood in terms of regular use, the previous explanatory failure queries the adequacy not merely of his theory of meaning, but also of his theory of truth. If the argumentation that I advanced in this section is correct, then we can establish that Horwich's minimal theory cannot explain all observable facts about truth. In particular, it cannot account for the objectivity of truth.

Finally, we may also note that if we adopt the above realist construal of truth, then Horwich's claim about the explanatory autonomy of our theory of truth with respect to our metaphysical, epistemological and other sorts of beliefs proves to be untenable too. Someone, for instance, who, on epistemological considerations, refuses to believe in the existence of platonic entities may have to implement some modifications in her theory of truth as well. As we have seen in chapter 3, in the semantics of discourses about abstract domains, the received realist, referentialist and non-revisionist construal of truth entails that in absence of platonic entities our widely shared beliefs about these

domains cannot be true. So, if an anti-platonist accepts that a proper theory of truth must be realist in character, and she is also convinced that a large part of our widely shared beliefs in pure logic and mathematics is true, then she must concede that at least one of her remaining two assumptions in the semantics of these discourses has to be modified: she must either give up the idea that the subject matter of pure logic and mathematics is abstract (i.e. non-spatiotemporal) in character, or replace the standard referentialist construal of these truths with a non-referentialist alternative. Either way, her original theory has to be modified, because it cannot account for at least some beliefs in metaphysics and epistemology.

Summing up, in this section, I examined the most fully elaborated version of deflationism in present-day philosophy, Paul Horwich's minimalist conception of truth, and argued that if we adopt, as I think we should, a use theory of meaning that guarantees the objectivity of correct use, then we can derive that Horwich's positive tenet about the notion of truth, namely that it can be adequately characterised by the instances of the equivalence schema, undermines his negative, deflationist claim that truth is not a substantive property whose nature and relation to our epistemic capacities could be further specified. In other terms, what I have shown is that if we adopt the above premises, then the positive tenets of the minimal theory entail that truth must be understood as a substantive real property, the property of having suitable semantic relations to some conditions that obtain in the actual world. Since Horwich explicitly rejects this realist construal of truth, he must also reject the corresponding realist construal of correct use. In absence of these realist tenets, however, his minimal theory will not be able to explain all facts about truth (and correct declarative use in general). In particular, it will not be able to account for the objectivity of this semantic property, the second *explanandum* specified in chapter 2. Finally, we observed that the realist perspective advocated here undermines the explanatory autonomy of the resulting theory of truth. If truth is understood as a real feature of our representations, then our beliefs about, among others, what there

is and how we discover what there is will impose substantial constraints upon our conception of the nature of this semantic property as well.

The core insight behind the above case against Horwich's minimalism, namely that without a substantive realist construal of the right-hand side of the instances of the equivalence/disquotation schema a theory of truth consisting of these instances cannot explain the objectivity of this semantic property, can be invoked in an argument against the explanatory adequacy of any form of deflationism. The specific feature of Horwich's minimalism is that it nicely illuminates the conceptual relation between our theory of truth and theory of meaning, and thus provides us with the necessary premises for deriving our case from basic intuitions about correct use in general, rather than merely about truth in particular.

A common feature of Horwich's minimalist and the suggested realist conception of truth is that they both understand truth in terms of the obtaining of some semantically significant conditions. The difference between the two conceptions is in the interpretation they provide of these conditions (or, rather, of the metaphysical/explanatory impact of their obtaining). According to Horwich, there are no explanatory considerations that would force us to adopt a substantive realist interpretation of the truth conditions whose obtaining is implied by the endorsement of the right-hand side of the instances of the equivalence schema. Thus he can maintain a referentialist construal of these conditions without thereby subscribing to controversial doctrines in metaphysics and epistemology.[14] In contrast, his realist opponent maintains that without a realist construal of these conditions the

---

[14] Reading Horwich, I have the feeling that one major motivation behind his efforts to develop a minimal theory of truth is the desire to escape the anti-naturalist consequences of traditional (i.e. referentialist) realism about truth in the semantics of our discourses about causally inert domains, a concern fully supported by the current work as well. What Horwich may intend to achieve by the denial of substantive realism about truth, I suggest achieving by the adoption of a non-referentialist form of realism in the semantics of our discourses about causally inert domains.

account cannot be explanatorily adequate, so the controversial doctrines in question must be avoided in some other way: either by insisting that, contrary to appearances, the conditions referred to on the right-hand side of the instances of the equivalence schema obtain within an unproblematic domain, or by conceding that the conditions whose (thick) obtaining is implied by the endorsement of this side are not necessarily the ones we refer to while making this endorsement.[15]

The idea that our semantically distinguishable beliefs have (mostly) different truth conditions enables us to develop a rather fine-grained explanatory apparatus to account for various truth-related phenomena in the world. Suppose, for instance, that I successfully found my glasses on the fridge, and also successfully made a 7 no-trump contract in a bridge game I played this evening. A relatively coarse-grained explanation of these achievements would refer, in both cases, to the fact that my beliefs on which my successful behaviour was based were true. A realist might add, what the minimalist would deny, that the same fact could have been identified by saying that the beliefs in question had suitable semantic relations to the actual world. Nevertheless, this supplement would not alter the fact that the two *explananda* have been explained by reference to the same *explanans*. Now, if our semantically distinguishable beliefs have, indeed, different truth conditions, then my achievements can be

explained in a more fine-grained manner as well, which corresponds to the intuition that the explanatory grounds of my success in the search, on the one hand, and in the game, on the other, are not entirely the same. According to these fine-grained accounts, the fact that my glasses were lying on the fridge (partly) explains my success in the search, but not in the game, while the fact that the dummy provided me with sufficiently high cards (partly) explains my success in the game, but not in the search. Clearly, if our semantically distinguishable representations did not have different truth conditions, then such sophistications in our understanding of these truth-related phenomena could not be achieved.

After presenting my argument against the deflationist construals of the notion of truth, in the following section, I shall turn to a case which purports to show that the received realist intuition that our semantically distinguishable true beliefs can refer to, and be made true by, different facts in the world is mistaken. Although the conclusion of the argument is compatible with a degenerated form of the suggested realist construal of truth, nonetheless it provides a serious challenge to the explanatory aspirations of any substantive theory of this semantic property, and thus can be read as an indirect case for the adequacy of deflationism.

2. Slingshot: A Case against the Plurality of Facts

A seemingly powerful objection to the received realist assumption that our true beliefs can refer to, and made true by, different facts in the world is to argue that (i) facts cannot be individuated independently of those representations to which they are supposed to correspond, and (ii) if they are individuated as the referents or truthmakers of particular representations, then we have no means to distinguish them from each other. Clearly, this result, if accepted, is devastating to any substantive (realist or anti-realist) theory of truth which purports to specify the nature of truth in terms of facts, the obtaining of some semantically

---

[15] If the right-hand side of the instances of the equivalence schema can specify the truth conditions of the propositions in the left-hand side in a non-referentialist way (i.e. without being about these conditions), as I think it must in the case of knowable propositions about abstract domains, then there must be pairs of beliefs with different semantic contents but identical truth conditions. This is because for each belief with non-referential truth conditions we will find a semantically distinguishable belief with referential truth conditions *about* the obtaining of those conditions. So, the fact that our belief *that p* and our belief *that q* have different semantic contents does not *guarantee* that the truth conditions of p and q are different too. Semantic contents are more fine-grained than truth conditions. In the case of an arbitrary truth-apt representation p, semantic contents include the declarative use conditions of p in larger (i.e. embedding) representational contexts as well, while the truth conditions of p are the declarative use conditions of p in an unembedded state.

significant conditions. An influential argument, commonly known as the "slingshot", has been advanced by a number of philosophers to establish from a few plausible tenets about the representational features of our thought and language the surprising conclusion that all true beliefs or sentences refer to the same thing.[16] If this is so, however, then the "thing" that a true belief is supposed to refer to cannot fulfil the explanatory role that the advocates of substantive theories of truth have attributed to it. Among other things, it cannot explain the intuitive difference between the (referential) truth conditions of various true beliefs. If all true beliefs are made true by the same Great Fact, then it is hard to see why it seems to us that these mental representations have different semantic contents, and their truth can be explained by reference to different aspects of the world.

For the purpose of the current discussion, in what follows, I shall briefly present the Gödelian form of the argument, whose assumptions are less contentious than those employed by the advocates of other formulations. Having reconstructed the case, I shall examine its major presuppositions and explain why I think that it cannot be taken as a convincing point against the realist (or anti-realist) belief in the existence of distinct referential truth conditions. For greater clarity, I shall present the argument using concrete examples. Suppose that the following three sentences are true in the actual world:

1. Peter is not identical with Thomas.
2. Peter is tall.
3. Thomas is short.

Beyond these factual presuppositions, the argument also assumes that the following three semantic principles hold:

I.   *Fa* and *a = (ιx)(x = a & Fx)* stand for the same fact.[17]
II. Any sentence that stands for a fact can be put into a predicate argument form.[18]
III. The referent of a composite expression, containing constituents which themselves have a referent, depends only on the referents of these constituents, and not on the manner in which this referent is expressed.[19]

From these presuppositions, Gödel's slingshot establishes that, contrary to appearances, (2) and (3) stand for the same fact. To see this, consider that, due to the specified assumptions, the consecutive members of the following series of sentences are all coreferential:

i.   Peter is tall.
ii. [Peter] is identical with [the object that is (tall and identical with Peter)].
iii. [The object that is (tall and identical with Peter)] is identical with [the object that is (not identical with Thomas and identical with Peter)].
iv. [The object that is (not identical with Thomas and identical with Peter)] is identical with [Peter].
v.   [Peter] is not identical with [Thomas]. // [Thomas] is not identical with [Peter].

---

[16] It is sometimes assumed that a version of this argument was behind Frege's famous decision to take all true sentences as referring to the value TRUE, while the false ones to the value FALSE. Frege (1892). Alternative formulations of the argument appear, among others, in Gödel (1944), Church (1943), Church (1956), Quine (1953b), Quine (1960b) and Davidson (1969). For a detailed presentation and analysis of these formulations, see Neale (1995), Neale (2001).

[17] The definite description '*(ιx)φ*' is supposed to stand for the unique thing which satisfies *φ*.
[18] Among others, this premise guarantees that an identity claim involving two singular terms ('*a = b*') can be equivalently regarded as the predication of identity of any of the denoted objects with the other (*IdB(a)* or *IdA(b)*). According to this premise all these representations are supposed to stand for the same fact.
[19] In other terms, the substitution of coreferential component expressions is not supposed to change the referent of the composite phrase.

vi.   [The object that is (not identical with Peter and identical with Thomas)] is identical with [Thomas].

vii. [The object that is (short and identical with Thomas)] is identical with [the object that is (not identical with Peter and identical with Thomas)].

viii. [Thomas] is identical with [the object that is (short and identical with Thomas)].

ix.   Thomas is short.

In particular, (i) and (ii), as well as (viii) and (ix), are coreferential because of (I); (iv), (v) and (vi) are coreferential because of (II); (ii), (iii) and (iv), as well as (vi), (vii) and (viii), are coreferential because of (III) and the coreferentiality of the relevant names and definite descriptions, which is supposed to hold because of (ii) and (iv), on the one hand, and (vi) and (viii), on the other.

The most popular way, among the friends of facts, to avoid the above collapsing conclusion is to adopt Russell's Theory of Definite Descriptions, according to which definite descriptions do not stand for objects as proper names do.[20] By this move, the derivation can be blocked, since the arguments of the main connective 'is identical with' in (ii), (iii), (iv), (v), (vi), (vii) and (viii) will no longer be coreferential, and therefore (III) will not guarantee the coreferentiality of the corresponding sentences either. From a semantical point of view, however, Russell's idea that definite descriptions cannot have (Fregean) referents in themselves, and thus cannot be coreferential with proper names, is at least as counter-intuitive as the anti-factualist conclusion that the move purported to neutralise.[21]

Another response to the argument could draw on the point suggested earlier that claims about relations among truth

---

[20] Russell (1905).

[21] In a world in which Peter is tall, the definite description 'the object that is tall and identical with Peter' can refer to a particular aspect of the world just as much as any proper name. As I shall argue, the problem with the slingshot is not that it presupposes this referential relation, but rather that it also assumes that the aspect associated with this definite description in the specified world cannot be distinguished from that associated with the proper name 'Peter'.

conditions should not be conflated with claims about relations among referents. If the epistemological case against the platonist construal of truth in the semantics of discourses about abstract domains is correct, then we have good reason to suppose that the truth conditions of these purportedly *a priori* commitments are different from the states of affairs that they intend to be about. Consequently, the conclusion of the slingshot should not be taken to establish anything about the truth conditions of these beliefs. One problem with this reaction is that it neutralises the slingshot only in the case of those beliefs whose truth conditions are non-referential in character. In the case of our broadly physicalistic claims, however, the adequacy of standard referentialism about truth has not been ruled out by the rejection of semantical platonism. Consequently, one might insist that the argument still demonstrates that at least in the case of our discourses *about the spatiotemporal world* all true beliefs are made true by the same thing. The second problem with the above reaction is that the slingshot's conclusion is counter-intuitive independently of what we believe about the relation of truth and reference. The surprising thing in this conclusion is that it contradicts our fundamental conviction that our true beliefs can have referential relations to different aspects of the world.

In my view, the real problem with the slingshot is that its presuppositions are informed by the Fregean coarse-grained idea of reference, which does not grasp those fine-grained referential aspects of our language which lend support to a referentialist understanding of truth conditions. The point will be clearer after examining the semantical assumptions (I) and (III) from the suggested referentialist point of view. Let me start with the latter. According to this assumption, for instance, the sentences (ii), (iii) and (iv) are coreferential. Why? Because all three are meant to refer to, and *prima facie* even made true by, the self-identity of the very same object, which is alternatively denoted by 'Peter', 'the object that is tall and identical with Peter', and 'the object that is not identical with Thomas and identical with Peter'. Although we can maintain that the coreferentiality of the three sentences relies on the obtaining of two further facts, namely Peter's being tall,

on the one hand, and his not being identical with Thomas, on the other, nevertheless, the three identity claims do not *refer to* those further facts, at least not in the received Fregean sense of the term. Frege's notion is intuitive enough. Of course, there is a sense in which the three identity claims are, indeed, *about* the same thing, namely the self-identity of the denoted object. This sense, however, is too coarse-grained to support a referentialist construal of truth conditions. This is manifest in the linguistic fact that while we may accept the claim that the three identity statements refer to the same fact, and at first glance maybe also that they are made true by the same fact, nonetheless, after realising the significance of those two further facts to the truth of (ii), (iii) and (iv), we become certainly more reluctant to subscribe to the second claim, independently of our view about the first. We will probably move toward the opinion that the truth conditions of (ii) involve the state of affairs that Peter is tall, those of (iv) that Peter is not identical with Thomas, while those of (iii) both of these states of affairs. Again, what this may show us is that our idea of truth conditions is more fine-grained than the Fregean concept of reference.

This result would already be enough to defend the classical realist view that our true beliefs possess their truth value in virtue of the obtaining of various states of affairs from the challenge of the argument. On the other hand, departing the field with this conclusion would leave our concept of reference, together with our notion of facts that we are supposed to think of or speak about, prey to the slingshot. The situation, however, does not seem to be even that bad. For there is clearly a sense in which (ii), (iii) and (iv) are not about the same aspect of the world. While (ii) is about Peter's being identical with the object that is tall and identical with him, (iv) is about Peter's being identical with the object that is not identical with Thomas and is identical with him. To be sure, in this fine-grained sense of the term, reference would embrace everything that Frege wanted to separate under the notion of sense. The fact that his terminology had well-known motivations in linguistic phenomena does not imply, however, that his choice optimally observed all theoretical

concerns. If we want to use the notion of reference to denote a substantive and sufficiently fine-grained semantic relation between our beliefs or sentences, on the one hand, and some particular aspects of the world, on the other, then we may well have to reconsider our received Fregean terminological heritage. Adopting the suggested fine-grained concept of reference, our principle of compositionality would still hold in the following, slightly modified form:

> III' The referent of a composite expression, containing constituents which themselves have a referent, depends on the referents of these constituents.

After abandoning the distinction between sense and reference, the clause 'and not on the manner in which this referent is expressed' at the end of (III) will lose its significance.

Turning to the first semantical assumption of the argument, we may draw more critical conclusions. Here again, we might admit that, in the Fregean coarse sense of 'reference', the sentences 'Peter is tall' and 'Peter is identical with the object that is tall and identical with Peter' refer to the same thing. Why? Because both are supposed to be made true by the same fact, namely Peter's being tall. In this case, to reply that 'because both are about Peter's being tall' is even less intuitive than the corresponding answer to explain the coreferentiality of (ii), (iii) and (iv) above. Be this as it may, we can certainly identify a sense in which (i) and (ii) are not about the same aspect of the world either. While (i) is about Peter's being tall, (ii) is about Peter's being identical with the object that is tall and identical with him. Adopting, again, this fine-grained concept of reference, we should conclude that the first assumption of the argument is simply false. Moreover, this construal would not rule out the referentialist understanding of truth conditions either. For, arguably enough, the truth conditions of (i) and (ii) are just as different as they were in the case of (ii), (iii) and (iv). The clearest evidence of this difference is the fact that if Peter were not tall, then (i) would be false (on *any* interpretation), while (ii) would be

either false or neither true nor false (depending on the semantics of the connective 'is identical with' and the definite description 'the object that is tall and is identical with Peter').

Summing up, in this section, I have argued that the case presented against the idea of separate facts, subject matters and truth conditions, and thus against any substantive theory that purports to specify the nature of truth in terms of distinct subject matters, relies on an intuitive, but rather coarse-grained idea of reference, which does not grasp those fine-grained referential aspects of our language that may lend support to a referentialist understanding of truth conditions in the semantics of broadly physicalistic discourses. If we realise that the standard referentialist conception of truth is based on a fine-grained notion of subject matter, which preserves the intuitive difference between the referential aspects of, say, (i) and (ii) or (ii) and (iii) and (iv), then we also understand why the slingshot provides us with such a surprising conclusion. It does so because it invites us to ignore a substantial part of those compositionally determined semantic relations that our truth-apt mental and physical representations have to various aspects of the world. If we refuse this invitation, then the slingshot collapses, and the idea of correspondence between truthbearers and particular aspects of reality can be maintained without inconsistency.

Summary

In this chapter, I examined the viability of the deflationist version of standard referentialism in the semantics of our cognitive discourses in general, and argued that the advantages the theory has (including the capacity to neutralise Benacerraf's challenge without abandoning referentialism in the semantics or inventing dubious explanations in the epistemology of pure logic and mathematics) are cancelled out by the fact that a deflationist can provide no suitable explanation of the objectivity of truth.

In section 1, I attempted to demonstrate this point by examining the most fully elaborated form of deflationism, Paul

Horwich's minimal theory of truth. Beyond providing a case for the claim that a proper account of this *explanandum* requires a substantive realist construal of truth, I have also shown that Horwich's positive tenets about meaning and truth entail the adequacy of a version of the standard realist correspondence theory of truth.

In section 2, I presented an argument which has been used to challenge an important presupposition of this realist account, namely that our semantically distinguishable true beliefs can refer to, and be made true by, different facts in the world. Having examined the major assumptions of the argument, I observed that at least two of them rely on a rather coarse-grained notion of reference, which cannot grasp those highly differentiated referential aspects of our thought and language that may lend support to a referentialist understanding of truth conditions in the semantics of broadly physicalistic discourses. The conclusion I drew was that with the adoption of a sufficiently fine-grained notion, we can easily block the slingshot and save the integrity of the suggested realist position.

In the following chapter, I shall turn to the second referentialist response to Benacerraf's dilemma left standing in chapter 3, which opposes the above realist account on epistemological considerations from the perspective of a substantive anti-realist theory of truth.