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

Novák, Z.

CHAPTER 6

Referentialist Responses to Benacerraf's Dilemma III: Platonist Construals of A Priori Truth and Knowledge

Introduction

In the previous two chapters, I argued that in order to explain the objectivity of our apparently legitimate knowledge claims, we need a substantive theory of truth whose subject matter is understood along realist lines. Since our purportedly *a priori* judgements about abstract domains seem to be as objective as our ordinary empirical beliefs about the spatiotemporal world, the above result is supposed to hold across the board in the semantics of all cognitive discourses, whether *a priori* or empirical. In chapter 3, however, we saw that in the standard referentialist framework a realist construal of truth in the semantics of claims about causally inert subject matters is incompatible with a causal theory of how we can acquire knowledge or reliable beliefs about such domains. So, if the explanatory considerations in support of realism about truth advanced in chapter 4 are correct, and consequently neither the deflationist nor the anti-realist forms of referentialism can provide a suitable response to Benacerraf's updated and generalised challenge in the philosophy of discourses about causally inert domains, then we must either abandon the standard referentialist framework and develop a non-referentialist construal of the problematic truths, or come up with a new account of how we are supposed to acquire knowledge or reliable beliefs about the obtaining or absence of causally inert conditions in the world.

In this chapter, I shall examine the prospects of the latter strategy in the semantics of our paradigm *a priori* discourses, such as logic and mathematics. The primary purpose of this investigation is to show that the advocates of platonism about the relevant truths have no suitable answer to Benacerraf's epistemological challenge presented in chapter 3, and that the responses they have actually given to this challenge are either *ad hoc* and uninformative, undermining all constructive methods for examining the nature and shortcomings of the relevant forms of belief formation, or insufficient, leaving us without any positive reason for supposing that the conditions whose obtaining they assume to be necessary and sufficient for the truth of our logical and mathematical beliefs indeed obtain in the intended platonic realms.

In section 1, I shall briefly review the most important explanatory considerations that may be raised for and against the platonist construal of our paradigm *a priori* truths. In section 2, I shall first examine the available platonist responses to what I regard as the most fundamental objection to this doctrine, namely Benacerraf's challenge that a platonist referentialist has no suitable account of how we could develop knowledge or reliable beliefs about platonic objects and properties, and then explain why I think that none of these responses can save the adequacy of the platonist construal of truth in the semantics of our paradigm *a priori* discourses about abstract domains.

1. The Explanatory Virtues and Vices of Platonism about **A Priori** *Truth*

As a point of departure, let me recall the most important motivations behind the platonist understanding of our knowledge claims about abstract states of affairs. According to Benacerraf, as we saw in chapter 3, one virtue of this construal is its *homogeneity* with our standard referentialist account of truth in the semantics of broadly physicalistic discourses. The essential tenet of this account is that the truth conditions of our legitimate knowledge claims are those states of affairs that these claims purport to be about. Since in the case of our paradigm *a priori* discourses the intended referents of our sentences are arguably abstract states of

affairs, the standard realist construal of these truths is certainly that provided by the advocates of semantical platonism. So, the first advantage of the platonist construal is that it meets the first adequacy condition set for such an account in the second part of chapter 2.

Beyond its homogeneity with our standard referentialist semantics, the construal provides a realist picture of the truth conditions of our paradigm *a priori* beliefs, and thus possesses the necessary explanatory resources to account for the objectivity of these truths. In this manner it satisfies the second adequacy condition listed in chapter 2 as well.

In chapter 5, we saw that the two most influential "anti-realist" attacks on the idea that our thoughts and sentences may have determinate semantic relations with those (often verification-transcendent) states of affairs that they seem to be about are equally unsound. In view of this result, we may assume that the platonist construal of the truth conditions of our paradigm *a priori* beliefs can be supplemented with a suitable account of the emergence of determinate semantic relations between the relevant *a priori* beliefs, on the one hand, and their intended abstract subject matter or truth conditions, on the other. At the very end of chapter 5, I provided an outline of a (moderate) causal account of reference determination, which, if true, also shows how the platonist construal satisfies the third adequacy condition on our list in chapter 2.

In possession of this account, a platonist can also explain how our paradigm *a priori* beliefs can be about an abstract and infinite domain (cf. the eleventh *explanandum* in chapter 2), and how we can develop, in principle, infinitely many semantically different representations of these domains (cf. the seventh *explanandum* in chapter 2). Her accounts can draw on the compositional character of semantic content. As to the former *explanandum*, one can think of an abstract domain by thinking of a domain that is not spatiotemporal, and one can think of an infinite domain by thinking of a domain that is not finite in character. The elements of this domain can be distinguished by invoking those (infinitely many) essential properties that can be recursively composed of the atomic properties of the domain. As to the latter *explanandum*, since the platonist understands the semantic content of our claims about abstract domains along the referentialist lines, she can explain the infinity of semantically different truth-apt representations within our paradigm *a priori* discourses by invoking the infinity of the available abstract truth conditions within the intended abstract domains.

Supposing that the abstract conditions thus related to our paradigm *a priori* beliefs are identical with the truth conditions of these truth-apt representations, semantical platonists can considerably enhance the apparent explanatory adequacy of their theory. By stipulating the necessity of the obtaining (or the absence) of the intended abstract conditions, for instance, they seem to be able to explain the necessary truth or necessary falsity of our paradigm *a priori* beliefs (cf. the ninth *explanandum* in chapter 2). Further, supposing that we have some sort of epistemic access to the intended non-spatiotemporal realms, the advocates of the platonist construal may offer a simple and natural account of the apriority of the evidential grounds or ways of justification that we rely on during the formation of these beliefs (cf. the eighth *explanandum* in chapter 2). Finally, by assuming the previous access, they can also explain the intersubjectivity of the relevant semantic contents and the observable convergence of the relevant beliefs (cf. the fifth and sixth *explananda*, respectively, in chapter 2).

In view of this remarkable explanatory potential, it is no wonder that platonist construals of the truth conditions of our paradigm *a priori* beliefs still preserve their appeal in present-day analytic philosophy. Nevertheless, as it is often emphasised, there are at least two major difficulties with this theory in the semantics of discourses about abstract states of affairs. The first is that the categorical separation of the domain of pure mathematical and logical beliefs from that of our empirical claims about the spatiotemporal world makes it hard to explain the applicability of the former types of knowledge claims in the empirical sciences

(cf. the tenth *explanandum* in chapter 2).[1] The second, explicated so aptly by Benacerraf and recast later in slightly different terms by Field, is that if the truth conditions of our logical and mathematical beliefs have no causal interaction with, or no influence on, the natural world, then the obtaining or absence of these conditions cannot be detected by spatiotemporally located human minds, which means that the platonist theory undermines the possibility, and the proper explanation, of mathematical and logical knowledge or reliable belief formation (cf. the third *explanandum* in chapter 2). In absence of a coherent notion of epistemic access to platonic realms, one may query the adequacy of the advanced platonist explanations of apriority, intersubjectivity and convergence as well, which considerably reduces the explanatory power of the platonist construal.

Of course, many platonists are aware of the significance of the previous explanatory difficulties, and they have various proposals of how these problems can be properly dealt with in the suggested platonist theoretical framework. In the following section, I shall examine the most influential platonist theories of how we can acquire knowledge or reliable beliefs about the posited platonic domains. I shall grant that the viability of any of these accounts would provide a decisive case for the adequacy of platonism about truth in the semantics of our paradigm *a priori* discourses, and thus support the cogency of standard referentialism as a conception of truth in general. The main message of the section, however, will be that the explanatory problem specified by Benacerraf's updated and generalised challenge presented in chapter 3 cannot be suitably resolved by any of the advanced platonist epistemologies. In view of this result, I shall conclude that the correct realist response to Benacerraf's challenge in the semantics of our paradigm *a priori* discourses (and our cognitive discourses about causally inert conditions in general) is to abandon the received referentialist construal of these truths and adopt a naturalistic understanding of

[1] For a detailed study of the problem of applicability of mathematics in our empirical theories of the world see Steiner (1998).

them that is compatible with a suitable account of all the *explananda* listed in chapter 2, and thus satisfies all adequacy conditions identified there for a proper theory of *a priori* truth.[2]

2. Platonist Accounts of Knowledge of Abstract Domains

Adopting Mark Balaguer's classification, platonist replies to Benacerraf's challenge can be grouped into two major categories.[3] One group maintains that, contrary to Benacerraf's assumption, human minds are capable of developing an epistemic access to platonic entities, while the other believes that knowledge of abstract domains does not require the mind's interaction with the obtaining platonic truth conditions. Following Balaguer's terminology, I shall call the two sorts of accounts, respectively, "contact theories" and "no-contact theories" of our knowledge of platonic states of affairs.[4]

[2] Since I believe that the epistemological challenge cannot be properly answered in the suggested platonist framework, I shall take it that the advocates of this construal have no suitable account for the apriority, intersubjectivity and observable convergence of the relevant beliefs either. Further, I believe that a platonist conception of truth in the semantics of pure logic and mathematics also undermines the proper understanding of how these purportedly *a priori* theories can be applied in our empirical accounts of the spatiotemporal world. In this chapter, however, I shall not argue for these further negative claims. I take it that the case provided here against the available platonist epistemologies is sufficient to show the inadequacy of the platonist construal. On the other hand, in chapter 7, I shall show that the suggested non-referentialist conception resolves all explanatory puzzles surrounding the relevant truths, including those left unexplained in the platonist semantical framework.

[3] Balaguer (1998), 24-25.

[4] In the philosophy of mathematics, Gödel's sporadic remarks on mathematical intuition are examples of the first theoretical alternative (Gödel (1944), 449, Gödel (1951), 310-312, Gödel (1964), 483-484). His views have been recently adopted by Brown (1991). BonJour's account of *a priori* knowledge provides a further, though slightly different, example of the first category. BonJour (1998). Authors falling into this contact theorist camp attempt to block Benacerraf's epistemological challenge by rejecting the first premise of his original argument reconstructed in chapter 3 (viz. that human beings exist entirely within space-time). Balaguer argues that the account developed by Maddy (1980) and Maddy (1990), according to which mathematical objects are spatiotemporal and human beings can acquire knowledge of them via sense perception, is also a contact theorist response to Benacerraf's case, although it queries the second, rather than the first, premise of the argument (viz. that if there exist any abstract mathematical objects and properties, then they exist outside space-time). In this chapter, I shall ignore the latter response, first, because it does not understand platonism in the traditional sense of the term, and second, because Maddy (1997) has abandoned this conception after all. Major examples of no-contact theories include Quine (1951), Steiner (1975), Parsons (1980), Parsons (1994), Katz (1981), Katz (1995), Resnik (1982), Resnik (1997), Wright (1983), Lewis (1986), Hale (1987), Shapiro (1989), Shapiro (1997) and Balaguer (1998).

The main problem with the contact theorist solutions is that their notion of a specific, non-causal information conveying channel between our minds and the alleged platonic realm is *ad hoc* and exotic in character. It is *ad hoc*, because by positing this epistemic link we can gain an account merely of the problematic *explananda* (knowledge acquisition and therewith apriority, intersubjectivity and convergence), and it is exotic, because we have no idea (neither *a priori* nor empirical) of the nature and working of this epistemic link in the actual world.

To appreciate this point, consider our reasons for adopting a contact theory of knowledge acquisition of the spatiotemporal world. By positing an epistemic contact (an information-conveying causal mechanism) between human minds and their natural environment, we can explain not merely the possibility of human knowledge of this domain (with the aposteriority, intersubjectivity and observable convergence of the relevant beliefs), but also our actual experience of the posited contact or mechanism.[5] In view of this extra explanatory impact, we can clearly reject the charge that our belief in this contact is *ad hoc* in character. Further, our theory of perception is an articulated, evidence-governed account of how the obtaining of natural states of affairs may influence our experience. It provides us with a

[5] Evaluating Putnam's just more theory argument against causal theories of reference determination, in chapter 5, I set forth which aspects of our experience can be taken as caused by the obtaining of the posited contact between our minds and their natural environment.

detailed picture of the nature of knowledge acquisition, and informs us about how we could eliminate our epistemic mistakes and improve the accuracy of our belief formation. To put it shortly, our idea of this contact is not exotic either. We can say that the observability of the posited epistemic contact is fairly expectable and also highly significant from the perspective of our cognitive purposes. On the one hand, it seems quite natural to suppose that if we can detect what obtains in some part of the mind-independent world, then in principle we must be able to detect the actual exercising of this epistemic capacity (in the same part of the world) as well. On the other hand, the observability of this contact is vital for both the improvement of our cognitive performance and our capacity to distinguish genuine knowledge from those Gettier cases in which our beliefs happen to be true without being properly informed by their obtaining truth conditions.

In contrast, the nature of our epistemic contact with the allegedly obtaining platonic truth conditions of our pure logical and mathematical beliefs seems fully inscrutable. We know that it cannot be causal in character. We are supposed to know this on *a priori* grounds: platonic entities cannot enter into causal relations. A platonist may add that the acquisition of this knowledge is also due to the existence of some epistemic contact between our minds and the obtaining truth conditions of this claim.[6] The assumption that our knowledge of the relevant contact is also due to the existence of such a contact does not create vicious circularity in the argumentation. Obviously, our knowledge of the nature of human perception (i.e. the epistemic contact between human minds and their natural environment) also presupposes the existence of the described perceptual links between our minds and the relevant perceptual relations in the world. The real problem with the platonist conception is rather that we can detect neither the actual obtaining (or absence) nor the existing

[6] One may wonder what a platonist would say about the nature of these truth conditions, as they involve the existence of a certain type of relation *between* the distinguished platonic and spatiotemporal realms.

characteristics of the posited epistemic relation in the actual world.

The lack of (*a priori* or *a posteriori*) observational evidence of the suggested epistemic contact between us and various platonic domains has also important theoretical and practical consequences. First, in absence of this observational ground we have no more reason for believing in the existence of the posited contact than a religious fundamentalist would have for her belief in the existence of a corresponding epistemic link between her mind and the allegedly obtaining truth conditions of her religious beliefs.[7] Second, in absence of that ground we cannot tell apart genuine knowledge from luckily acquired true beliefs either (i.e. we cannot tell whether a certain piece of *a priori* evidence in someone's mind is part of the posited epistemic link between that mind and the obtaining truth conditions of the relevant *a priori* beliefs, or rather it is created by some natural mechanisms that are entirely independent of what obtains in the posited platonic domains).[8] Finally, in absence of the relevant observations, we cannot develop an articulated theory of the nature of *a priori* knowledge acquisition, and we cannot learn how we could eliminate our epistemic mistakes and improve the accuracy of our belief formation in our discourses about abstract domains.

Summing up, contact theorist responses to Benacerraf's original or modified and generalised challenge to platonism about truth in the semantics of our discourses about causally inert

[7] The fact that the congruent belief-systems of various religious communities (or individuals) are incompatible with each other need not undermine the appeal of the suggested contact theory of religious knowledge acquisition. For someone who believes in that contact, it merely demonstrates that some of the conflicting alternatives must be mistaken and consequently the stipulated link does not provide us with absolutely reliable beliefs.

[8] One may think that the latter difficulty can be avoided by denying that beliefs about abstract domains are based on spatiotemporal evidence, and maintaining that knowledge within the relevant discourses consists in a direct grasp of the obtaining platonic truth conditions. Note, however, that this solution would undermine the distinction between justified and unjustified true beliefs about abstract domains, and thus contradict some of our basic intuitions in the epistemology of pure logic and mathematics.

domains are all inadequate, in so far as the explanation they provide of our knowledge or reliable belief formation about these domains is *ad hoc* and exotic in character. The adoption of this platonist epistemology would discourage any further inquiry into the nature of this type of knowledge acquisition, and it would open the door for parallel stipulations in the case of any other knowledge claims, no matter how the beliefs in question were causally produced in the subjects' minds.

The majority of contemporary platonists in philosophy of mathematics prefer the second type of account of how knowledge of abstract states of affairs is possible. The common feature of these accounts is the denial of the *prima facie* plausible claim that mathematical knowledge requires epistemic contact with the obtaining truth conditions of our correct mathematical beliefs. Instead, the advocates of these no-contact theories argue that some specific properties of these platonic truth conditions and/or the way we develop our beliefs about them in the spatiotemporal world guarantee and explain the possibility of mathematical knowledge or the reliability of mathematical belief formation.[9] Of course, the crucial question in this case is whether the invoked characteristics are indeed sufficient for ensuring the envisaged result. The bare stipulation of the possibility of mathematical knowledge acquisition or reliable belief formation in space and time would hardly satisfy those who share Benacerraf's reservations about the platonist construal of mathematical truth.[10]

[9] Authors falling into this no-contact theorist camp attempt to block Benacerraf's epistemological challenge by rejecting the third premise of his original argument reconstructed in chapter 3 (viz. that if there exist any abstract mathematical objects and properties, then human beings cannot have knowledge of them), while maintaining the first and the second about the nature of human beings and mathematical entities, respectively.

[10] As Field formulated, "special 'reliability relations' between the mathematical realm and the belief states of mathematicians seem altogether too much to swallow. It is rather as if someone claimed that his or her belief states about the daily happenings in a remote village in Nepal were nearly all disquotationally true, despite the absence of any mechanism to explain the correlation between those belief states and the happenings in the village". Field (1989), 26-27. Balaguer

One influential strategy to account for the possibility of mathematical knowledge, which may be regarded as a no-contact theory of knowledge of platonic facts, is to adopt Quine's holistic theory of confirmation, and argue that our mathematical beliefs constitute a (maybe indispensable) part of our overarching theory of the world, which as a whole is confirmed by the (naturalistically construable) deliverances of our external and internal senses, in other words by our experience.[11] Of course, this account can be taken as a no-contact theory of knowledge of platonic facts only if we suppose that the truth conditions of our mathematical beliefs obtain in a platonic realm. Whether Quine himself maintained a platonist construal of mathematical truth is rather questionable. Those who believe that he did may rely on his famous verdict on ontological commitment. According to this proposal, our prevailing idea of what there is is determined by our best overall theory of the world: in particular, we are committed to the existence of those entities that this theory happens to quantify over.[12] Since, in Quine's view, mathematics is an integral part of our best overall theory of the world, one may conclude that Quine must have believed in the existence of mathematical objects.[13] Note, however, that Quine is also famous

(1998) formulates the same point against Parsons's theory of mathematical knowledge, which purports to explain the phenomenon by invoking the epistemic capacity of intuiting "quasi-concrete" objects (i.e. types of perceivable tokens) without explaining why the deliverances of this capacity would provide reliable information of "purely abstract" (i.e. platonic) states of affairs. Balaguer (1998), 38, esp. footnote 46, and Parsons (1980).

[11] Balaguer (1998), 40-41. Advocates of this account of mathematical knowledge include Steiner (1975), Resnik (1997) and Colyvan (2001), Colyvan (2007). We may note also that Field's reading of Benacerraf's epistemological challenge, according to which the problem with the platonist construal is that it cannot explain the reliability of our actual mathematical beliefs, was largely put forward also as a reaction to this Quinean response to Benacerraf's original argument. Field (1989), 25.

[12] Quine (1948), Quine (1951).

[13] The same reconstruction of Quine's platonist reading appears in Hellman (1989), 3, fn. 1. In so far as the theory can be formulated, as Quine believed, in first-order language, the commitments in question will merely extend to the domain of the first-order variables.

for his deflationist (disquotational) theory of truth, which is hardly compatible with the substantive realist construal of the obtaining truth conditions of true mathematical beliefs implicated by the platonist construal.[14] Further, Quine's empiricist theory of confirmation was meant to provide an account of mathematical knowledge only in so far as mathematics is (indispensably) applied in our best overall scientific theory of the spatiotemporal world.[15] Since the truth conditions of this theory are supposed to be spatiotemporal, it is far from obvious how Quine's epistemology could qualify as a no-contact theory of knowledge of platonic domains.

Now, of course, independently of these interpretative questions, one may adopt the above empiricist strategy and maintain that our external and internal senses (i.e. our causal-epistemic contacts with the natural world) provide us with knowledge not merely of the spatiotemporal world, but also of the obtaining platonic truth conditions of our claims about abstract domains. In the philosophy of the relevant discourses, such as pure logic and mathematics, this account clearly qualifies as a no-contact theory of knowledge of platonic domains, since it preserves the idea that the truth conditions of these beliefs obtain in a platonic realm, and it does not presuppose the existence of any contact between the posited platonic objects and properties and our knowing minds. Instead of stipulating such a contact, the account rather explains the way we acquire knowledge of platonic entities by emphasising that this knowledge is inseparable from (and maybe indispensable to) our knowledge of the

[14] Quine (1970).
[15] Although Quine was quite hesitant about how much of our mathematical theories can be legitimately justified on holistic empirical considerations, he made it explicit several times that those parts that are demonstrably independent of the applicable pieces must be regarded as results of mathematical recreation, whose acceptance does not presuppose specific ontological commitments on our part. Quine (1986), 400, Quine (1995), 56-57. For a defence of this Quinean semantical division within mathematics see Colyvan (2007).

spatiotemporal world.[16] In fact, we can justify our beliefs in the existence of the relevant platonic objects and properties in the same way as we justify our beliefs in the existence of theoretical entities posited by our best overall theory of the spatiotemporal world: first, we justify our overall theory holistically, in view of its predictive success, on empirical considerations, and then, we understand its truth in referentialist terms (i.e. in terms of the existence of those entities and the obtaining of those conditions that the theory purports to be about). The question, of course, is whether we can indeed legitimately suppose such a parallelism between the two types of knowledge acquisition.

Opponents may query the correctness of this assumption on various considerations. They may observe, for instance, that platonic objects and properties, unlike electrons and their particular features in space and time, cannot be invoked in causal explanations of phenomena, so their existence cannot have the same explanatory role in our theory of knowledge as that of entities posited by our best overall theory of the spatiotemporal world. One may think that the contrast mentioned by these opponents wrongly presupposes that the causally relevant properties of electrons and other theoretical entities can be represented without reliance on concepts of abstract objects and properties.[17] But to think that they cannot seems to rest on the relatively entrenched and arguably mistaken view that our concepts of mathematical and logical properties appearing in space and time are dependent on (or posterior to) our concepts

[16] Today, it is often granted that the only serious consideration in support of the platonist belief in the existence of (non-spatiotemporal) mathematical objects is Quine and Putnam's indispensability argument. Classical formulations of the argument occur in Quine (1948), Quine (1960a), Putnam (1971), and Putnam (1975d).
[17] The point appears in Field (1989). His conclusion on the matter is that "unless a very substantial amount of explanation involving electrons can be given in a mathematical entity-free fashion, the prospects for maintaining realism about electrons without maintaining platonism are dim". Field (1989), 19-20. Earlier Field made an attempt to demonstrate that our scientific theory of the world does not rely essentially on ideas of platonic entities. Field (1980).

of platonic objects and properties in pure logic and mathematics. In chapter 7, I shall argue that the dependence between these types of concepts indeed obtains, but it holds the other way round: our ideas of abstract (i.e. non-spatiotemporal) entities in pure logic and mathematics are dependent on (or posterior to) our concepts of mathematical and logical properties appearing in space and time.[18] If so, however, then the previous observation of the categorical difference between the explanatory role of platonic objects and properties, on the one hand, and those of the theoretical entities of our empirical sciences, on the other, may be fully adequate. In view of the causal inertness of platonic objects and properties, our belief in their existence cannot be based on the same explanatory considerations as our belief in the existence of the theoretical entities of the spatiotemporal world.

Realising that our reasons for believing in the existence of theoretical entities (together with their logical and mathematical properties) in space and time cannot support our belief in the existence of those abstract objects and properties that constitute the subject matter of pure logic and mathematics may lead the (realist) opponents of the empiricist account in one of the following two directions: they may either accept confirmation holism (i.e. empiricism) in the epistemology of our discourses about abstract domains and reject the referentialist construal of truth (i.e. platonism) in the semantics of these sorts of claims, or they may reject confirmation holism and check whether there is a better (presumably apriorist) no-contact epistemology which can account for our knowledge or reliable belief formation in discourses about abstract domains in a referentialist semantical framework.[19] If they choose the second option and their answer to the previous question is negative, then they must conclude that a correct theory of knowledge of abstract domains requires the adoption of a non-referentialist construal of the relevant truths.

[18] For a proper explication of the often conflated notions of abstractness, see section 5 in chapter 1.

[19] I ignore here the realist reactions disqualified earlier in chapters 3, such as scepticism and revisionism about subject matter.

Either way, the idea of an empiricist no-contact theory of our knowledge of platonic domains is abandoned for the sake of some alternative.

Although the adoption of the first alternative (i.e. empiricism with non-referentialism) is outside the scope of the current chapter, it may be worth briefly reviewing the two main reasons for which this strategy is seen by many philosophers as a non-starter.

First, one may observe that our beliefs in pure logic and mathematics do not rely on the applicability of these theories in our best overall theory of the spatiotemporal world. The fact, for instance, that Euclidean geometry is strictly speaking no longer applied in our scientific theory of the physical world does not influence our beliefs about the objects and properties posited by this geometry. Apparently, once we develop our concepts of the relevant abstract subject matters, our reasons for adopting or rejecting a claim that involves some of these concepts will have nothing to do with our empirical findings about the spatiotemporal world. Of course, this is merely a restatement of the eighth major *explanandum* put forward in chapter 2: namely, that our beliefs about abstract domains are based on *a priori* evidence.[20] So, the first major objection to the empiricist account under discussion is that it runs counter the apparent apriority of our knowledge of the relevant abstract domains.

The second objection is strongly related to the first. It starts with the observation that the truth or falsity of our beliefs about abstract domains is necessary in character (cf. the ninth *explanandum* in chapter 2), and then it lays down that our experience can never provide us with reasons for believing that something is necessarily true or necessarily false. In other terms, the second major objection to the empiricist theory is that it fails

[20] To avoid empty terminological objections to the idea that our beliefs about abstract domains are *a priori* in character, see my clarificatory notes on apriority in the first section of chapter 1.

to account for our knowledge of the modal character of the truth value of our beliefs about abstract domains.[21]

In view of these explanatory problems with the empiricist account of our knowledge of abstract states of affairs, we may conclude that the more plausible move after the realisation of the inadequacy of the earlier "Quinean" strategy is to proceed in the second direction (i.e. to reject radical confirmation holism and check whether there is an apriorist no-contact epistemology which could account for our knowledge of abstract domains in a referentialist semantical framework).

One influential example of such no-contact theorist replies to Benacerraf's challenge has been advanced by Bob Hale and Crispin Wright.[22] According to their view, our ability to develop thoughts and acquire knowledge of platonic states of affairs cannot be problematic, as Benacerraf implies, since the truth conditions of these beliefs are, in fact, identical with the truth conditions of some other, semantically and epistemologically unproblematic, claims. The core idea behind this suggestion is that we acquire our concepts and knowledge of platonic entities by Fregean abstraction and deductive inference without ever being acquainted with (or influenced by) the entities themselves.

Consider, for instance, the case of our concepts and knowledge of natural numbers. According to Frege, the acquisition of numerical concepts requires the acquisition of the truth conditions of all identity statements involving these concepts. In Hale and Wright's reconstruction, the latter can be done by means of what has come to be called *Hume's Principle*:

(HP) The number of *F*s = the number of *G*s ↔ there is a one-one correspondence between the *F*s and the *G*s.[23]

[21] The objection plays a central role in Katz's argumentation against empiricism in the epistemology of discourses about abstract domains. Katz (1981), 208.
[22] Hale and Wright (2002).
[23] As Hale and Wright rightly observe, if the principle is meant to inform us about the truth conditions of the identity statements on the left hand side of the sentential connective '↔', then the connective must be taken to indicate that the truth conditions of the connected expressions are identical, rather than merely that the connected statements are materially equivalent. Hale and Wright (2002), 117. Note also that Frege did not think that we can contextually define our numerical concepts by means of Hume's Principle, since the latter does not appear to enable us to settle the truth value of identity claims that link denoting expressions involving numerical concepts of the form 'the number of …' with others not doing so, such as the claim "the number of Jupiter's moons = Julius Caesar". Frege (1884), 67-68. In view of this problem, Frege famously decided to identify cardinal numbers with extensions, which move rendered his system of arithmetic inconsistent, entailing Russell's paradox. Frege (1884), 79-80. For a detailed discussion and treatment of the *Caesar Problem* (in response to Dummett's criticism of the neo-Fregean account in Dummett (1991)), see Hale and Wright (2001b).

A similar principle, the so-called *Direction Equivalence*, is supposed to underlie our acquisition of the geometrical concept of abstract direction:

(DE) The direction of line *a* = the direction of line *b* ↔ lines *a* and *b* are parallel

As Hale and Wright observe, (HP) and (DE) are two instances of the same general *Abstraction Principle*:

(AP) $\forall\alpha\forall\beta\ (\Sigma(\alpha) = \Sigma(\beta) \leftrightarrow \alpha \approx \beta)$

where '$\approx$' stands for an equivalence relation on entities of the type of α, β, and Σ is a function from entities of that type to objects.[24] According to Hale and Wright, (AP) provides us with a general tool to formulate thoughts and acquire knowledge of platonic

[24] Hale and Wright (2002), 118. One may argue that the adoption of this principle in its full generality is certainly not admissible, since some of its instances, most notably Frege's Principle of Extensional Abstraction (Basic Law V of *Grundgesetze*), lead to a contradiction. In view of this problem, Dummett (1991) warned that in absence of an explicit specification of what distinguishes the harmless instances of (AP) from the harmful ones, the neo-Fregean reliance on (HP) and (DE) is cannot be justified. For Hale and Wright's response to this "bad company argument", see Hale (1994), and Wright (1998). For an in depth discussion of abstraction principles in general, see Fine (2002).

objects without having an epistemic contact with them. On the one hand, with the help of (AP)'s instances, we can implicitly define concepts referring to abstract objects, and with the ensuing conceptual apparatus we can form beliefs about platonic realms. On the other hand, the above derivation of mathematical concepts is a key to a no-contact theorist explanation of mathematical knowledge as well. According to this conception, mathematical knowledge is grounded on knowledge of the identity conditions of mathematical objects, which merely requires that we can discern whether the truth conditions of identity claims about those objects (i.e. instances of '$\Sigma(\alpha) = \Sigma(\beta)$' in (AP)) obtain. These conditions, however, are stipulated to be identical, by the instances of (AP), with the truth conditions of some epistemologically unproblematic claims of equivalence relations (i.e. instances of '$\alpha \approx \beta$' in (AP)). Clearly, if our knowledge of the instances of (AP) and the relevant equivalence relations is unproblematic, then our knowledge of the corresponding abstract states of affairs cannot be problematic either.[25]

As it might be expected from what has been said so far, there are a couple of things on which I agree with Hale and Wright. First, I agree that we can think of and speak about platonic entities, and that the intended referents of our pure mathematical beliefs are abstract in the relevant sense of the term. Second, I also agree that we can acquire knowledge of such abstract entities, and that mathematics is the collection of such knowledge. Finally, I agree even that our knowledge of

[25] As Hale and Wright puts it: "So long as we can ascertain that lines are parallel, or that concepts [in the Fregean sense, Zs. N.] are one-one correspondent, there need be no *further* problem about our knowledge of certain basic kinds of truths about directions and numbers, for all their abstractness. For provided that the concepts of *direction* and *number* can be implicitly defined by Fregean abstraction, we can know statements of direction- and numerical-identity to be true just by knowing the truth of the appropriate statements of parallelism among lines and one-one correspondence among concepts. We can do so for the unremarkable reason that the truth-conditions of the former are fixed by stipulation to coincide with those of the latter." Hale and Wright (2002), 119.

mathematical and other abstract entities does not require an epistemic contact between us and the constituents of a platonic realm. What I do not agree with is that the previous commitments entail the endorsement of a platonist construal of mathematical truth (or truths about abstract domains in general) combined with a no-contact theorist response to Benacerraf's epistemological challenge to this semantical position.

There are at least two reasons for querying the correctness of the neo-Fregean transition from the former premises to the latter conclusion. First, the subject matter of a thought or sentence can be abstract without being platonic in character. Directions, for instance, as universal properties can characterise fictive and real spatiotemporal objects as well.[26] An account of our ability to refer to and acquire knowledge of abstract entities may therefore amount to a platonist response to Benacerraf's challenge only if the subject matter of this knowledge is abstract in the required sense of the term. In contrast to the numerical terms of pure mathematics, however, the denoting expressions appearing on the left hand side of the instances of Frege's Abstraction Principle do not necessarily stand for abstract entities in the required sense of the term (i.e. they do not necessarily stand for strictly non-spatiotemporal entities). The expression 'the number of apples in front of me', for instance, primarily stands for a property of a group of objects in the spatiotemporal world, rather than for an object of a platonic realm.[27] Other

[26] In the fifth section of chapter 1, I mentioned that a proper understanding of the epistemological problem with platonist theories of truth requires the careful disambiguation of our notion of abstractness, and separated the abstractness of spatiotemporal properties from the abstractness of entities that are supposed to exist in a platonic realm.

[27] In fact, the expression can be used to refer to both a numerical property of a group of objects in the spatiotemporal world and a mathematical object outside space and time. Due to this ambiguity, the sentence 'the number of apples on this table is the same as the number of spoons' can be interpreted also in at least two ways. On the first reading, it expresses a synthetic proposition about the numerical properties of the apples and the spoons on the table, while on the second, it expresses an analytic proposition about some mathematical objects outside space and time. On the account of semantic content advocated in this

denoting phrases, such as the expression 'the number of primes between 70 and 80', refer to a property of non-spatiotemporal entities, but the contextual definitions provided by Hume's Principle in these examples presuppose that we already acquired some concepts and knowledge of a non-spatiotemporal domain. Putting it briefly, what the neo-Fregean account seems to explain is how we can develop new concepts of certain fields from some earlier acquired ones of the very same fields. What it fails to explain is how we develop our notions of entities that cannot appear in space and time in the first place, maybe relying on our notions acquired earlier of entities appearing in space and time.[28]

work, the ambiguity is a consequence of those referential intentions that lie behind the two sorts of applications. While in the first case we intend to make an empirical claim about the contingent numerical properties of two groups of objects in space and time, in the second scenario we intend to advance an *a priori* statement of the necessary self-identity of a mathematical object within the non-spatiotemporal domain of pure mathematics. (Thanks to András Simonyi for reminding me of this ambiguity.)

[28] I admit that by stipulating that Σ is a function from entities of the type of α, β etc. to objects (i.e. individuals) Hale and Wright ensure that the concepts resulting from Fregean abstraction stand for entities of the right kind. Note, however, that the stipulation entails that the expressions 'the number of apples in front of me' and 'the direction I am actually looking at' are about strictly non-spatiotemporal mathematical objects, rather than about some properties that may characterise entities in space and time. As the observation made in the previous footnote may show, the problem with these readings is not that they are inadequate in the light of our actual linguistic and cognitive practice. By the application of these expressions we can definitely talk about the strictly non-spatiotemporal individuals of pure mathematics as well. The problem with the stipulation is rather that it does not provide us with those readings which explain the plausibility of the neo-Fregean claim that the truth conditions of the expressions on the two sides of the sentential connective '↔' in (AP) are identical. The reason for which we accept this claim is, to adopt Frege's own formulation, invoked by Hale and Wright as well, that we believe that the expressions on the two sides of the instances of (AP) "carve up" the same content in different ways. Frege (1884), 75, Hale and Wright (2002), 117. What this metaphor suggests is that the expressions on the two sides of the instances of (AP) have the same truth conditions, because they state the obtaining of the same conditions in slightly different ways. Supposing that the unproblematic expressions on the right-hand side of the instances of (AP) are not about the strictly non-spatiotemporal domain of pure mathematics, we can derive that the representations on the left-hand side are not meant to be about that domain either. If we take the intuitive ground of the neo-Fregean claim about the identity of the relevant truth conditions seriously, then we cannot suppose that Σ in (AP) is a function from entities of the type of α, β etc. to individuals within a strictly non-spatiotemporal domain. Putting it briefly, the way in which Hale and Wright invoke (AP) in their reasoning against their anti-platonist opponents raises serious doubts about what they really mean when they stipulate that the concepts generated by Fregean abstraction denote *objects* of a certain kind. In chapter 7, I shall suggest that our concepts of strictly non-spatiotemporal entities are developed by a cognitive operation that might be formally characterised along the neo-Fregean lines, but will not support the related neo-Fregean response to Benacerraf's epistemological challenge to platonism about truth in the semantics of our paradigm *a priori* discourses about non-spatiotemporal domains.

So, the first reason for querying the neo-Fregean view that we can acquire concepts and knowledge of platonic entities by means of Fregean abstraction is that the notions defined contextually by the instances of Frege's Abstraction Principle are generated by a sort of abstraction that does not guarantee the atemporal character of the resulting intended referents.

Second, even if we grant, as I think we should, that by a certain sort of abstraction we can develop concepts of strictly non-spatiotemporal entities as well, an account of knowledge of such entities can amount to a platonist response to Benacerraf's challenge only if the truth conditions of the thoughts composed of these concepts are to be construed in a realist and referentialist way. Now, the question that I propose to consider is whether the neo-Fregean epistemology advocated by Hale and Wright is compatible with such a realist and referentialist construal of mathematical truth. In what follows, I shall show that the proper response to this question is also negative.

Let us forget for a moment the previous observation that the denoting expressions on the left hand side of the instances of Frege's Abstraction Principle do not necessarily stand for abstract entities in the required sense of the term. According to Hale and Wright, our knowledge of non-spatiotemporal objects, such as numbers and (platonic) directions, must be unproblematic, since this knowledge is logically derivable from those identity claims which stand on the left hand side of the instances of Frege's

Abstraction Principle, whose truth conditions in turn are identical with the truth conditions of those, epistemologically unproblematic, beliefs, which stand on the right hand side of the relevant instances of Frege's Abstraction Principle. Supposing that to be non-problematic, in the context of Benacerraf's epistemological challenge to platonism about mathematical truth, means being non-platonic in character, what the instances of Frege's principle suggest to us is, among others, that the obtaining truth conditions of our mathematical beliefs are not platonic in character. If this is true, however, then the referentialist construal of these conditions must be abandoned: if our mathematical beliefs are about strictly non-spatiotemporal entities, while their truth conditions are supposed to obtain in an epistemologically unproblematic (presumably spatiotemporal) realm, then the latter conditions cannot be construed in terms of the former subject matters. If they could, then, again, according to the neo-Fregean reasoning, the obtaining truth conditions of our beliefs about the relevant right-hand equivalences would be platonic as well, and our knowledge of the latter would be no less problematic from Benacerraf's perspective than our knowledge of other platonic objects and properties in general. So, the second reason for querying the neo-Fregean view that we can acquire knowledge of platonic entities by means of logic and Fregean abstraction is that a charitable interpretation of the neo-Fregean account is incompatible with a realist and referentialist (i.e. platonist) construal of mathematical truth.

Apparently, Hale and Wright do not realise this incompatibility. They believe that the (realistically construed) truth of a certain kind of belief (here, a belief about non-spatiotemporal objects and properties) implies the real existence of the intended subject matter. Consider the following summary of their position:

> [I]n order to establish an intelligible use for singular terms purporting reference to numbers, or other abstract objects – that is, objects which are not 'external' (located in space), and of which we can have no 'idea' or 'intuition', but which are, in Frege's view, nonetheless objective – it suffices merely to explain the truth-conditions of statements incorporating such terms. No precondition involving prior engagement with or attention to the referents of such terms is soundly imposed. Moreover if, under a suitable explanation of the truth-conditions of an appropriate range of such statements, suitable such statements are – or may warrantedly be claimed to be – true, then those of their ingredient terms which purport reference to numbers or other abstract objects will in fact refer – or may warrantedly be claimed to succeed in referring – to such objects; and the intelligent contemplation of such a statement will constitute thought directed upon the objects concerned.[29]

Is there a legitimate motivation behind the maintenance of the referentialist construal underlying these formulations? As a first reaction, one may argue that if Hale and Wright's conviction were true (i.e. the realistically construed truth of a certain belief indeed implied the existence of the intended subject matter of the belief), then we could not form such true beliefs about fictive entities. Creating a fiction consists of the stipulation of a number of facts in an invented (i.e. paradigmatically non-real) universe. So far as we are told, for instance, in the fictive universe of Little Red Riding Hood it is a stipulated fact that the girl has got a grandmother. The stipulation, as opposed to the girl and the grandmother, is a fact of the real world, which guarantees that the sentence 'Little Red Riding Hood has got a grandmother' is objectively true, independently of whether anyone ever recognises this circumstance.[30] So, we may assume that the above sentence

---

[29] Hale and Wright (2002), 115.

[30] One may, of course, start or continue a story and maybe even change the truth value of some claims within the narrative without thereby changing the intended subject matter of the applied component expressions. A claim during the actual creation of a fictive story has no realistically construable truth conditions. Its truth

expresses a (realistically) true belief about Little Red Riding Hood. Applying Hale and Wright's referentialist tenet to this case, one may argue we must conclude that our concept of Little Red Riding Hood "may warrantedly be claimed to succeed in referring" (i.e. that Little Red Riding Hood exists just as much as the abstract referents of our mathematical claims). This conclusion, however, would be obviously false, so Hale and Wright's conviction cannot be generally true either.

To this objection a neo-Fregean may answer that successful reference is not meant to imply real existence. It is merely meant to imply "intended sort of existence": real and abstract in the case of mathematical objects and properties, while fictive and mostly spatiotemporal in the case of fictive entities, such as Little Red Riding Hood. Clearly, there is an important difference between the semantic content of the contrasted types of beliefs. While our ideas of mathematical objects and properties do not exclude the metaphysically thick or "real" existence of these subject matters, our referential intentions in the course of thinking of a fictive domain imply that the intended subject matters "exist" only thinly, in the relevant fictive universe. Little Red Riding Hood cannot be part of the real world, mathematical objects and properties in principle can.[31]

The contrast is apparently correct. Contrary to the case of our thoughts of fictive entities, our referential intentions in pure logic and mathematics do not exclude the existence of the subject matter of our true beliefs. Allowing the existence of these subject matters, however, is not the same as guaranteeing or requiring

value is determined by the authors who stipulatively characterise the relevant fictive universe. Nevertheless, once the story acquires its canonic form, the stipulations made provide a factual base (in the actual spatiotemporal world) for the evaluation of any further claims about that universe.

[31] In case the neo-Fregean accepts that the truth conditions of our claims about fictive entities can be construed along realist lines, the current response would amount to the adoption of a non-referentialist account of truth in the semantics of discourses about fictive domains.

(for truth) their existence.[32] Hale and Wright's position, that the existence of mathematical knowledge of non-spatiotemporal entities presupposes the existence of these entities would be correct if our cognitive practice in pure mathematics involved the referentialist idea that our beliefs cannot be (realistically) true unless the state of affairs they are about obtain in the real world, rather than merely in a world invented and projected by human minds. Note, however, that this idea is not inherent in daily mathematical practice. It is merely the manifestation of a substantive metaphysical and semantical position in the philosophy of mathematics. We saw that this position has problematic implications in the epistemology of mathematics: its adoption undermines the possibility of mathematical knowledge. No-contact theorist accounts of knowledge of platonic domains, in particular, do not attribute a substantive explanatory role to the posited platonic objects and properties, and thus, as we saw it in the empiricist case, undermine the idea that we can have an epistemic ground for believing in the existence of these entities and the truths that this existence allegedly constitutes.33 Further, I argued that the platonist position is incompatible, under a charitable interpretation, with the suggested neo-Fregean account of mathematical knowledge and truth. Finally, we may add that the unrestricted approval of the referentialist principle, which sanctions the inference from (realistically construed) truth to the

[32] In chapter 7, I shall argue that no analytic truth requires the reality of the relevant intended referential domains. What the realist construal of these truths requires is merely the reality of the crucial link between the relevant conceptual constituents of the analytic claims under scrutiny.

[33] Here I am relying on the naturalistic conception of evidence advanced in chapter 2, according to which theories can be justified by reference to a certain pool of evidence if and only if, and because, by reference to their obtaining truth conditions we can explain the actual occurrence of this evidential ground. Since by referring to platonic entities we cannot explain the actual occurrence of anything in the spatiotemporal world, nothing in this world can be legitimately taken as an epistemic ground for adopting or rejecting a belief that has platonic truth conditions.

existence of subject matter, opens the gate for stipulating or defining entities into existence.[34]

So, why should a neo-Fregean insist on this referentialist (and thus platonist construal) of mathematical truth? If in the semantics of our discourses about abstract domains we abandon that referentialist conception, and adopt an epistemologically unproblematic naturalist construal of the truth conditions of our beliefs, then we can maintain the idea of objective knowledge of abstract entities while remaining agnostic about the actual existence of the intended abstract subject matters. The resulting position could embrace platonism about referents or intended subject matters, but it would imply anti-platonism about the relevant truths.[35] Anti-platonism does not imply anti-realism

[34] The same charge appears in Field (1984), Dummett (1991) and more recently in Potter and Smiley (2001). Hale and Wright's response to the charge is that the instances of Frege's Abstraction Principle do not stipulate into existence any objects. What they do create is merely a certain sortal concept. As Hale observes, "[w]hether that concept is instantiated is, always, a matter settled by the truth or falsity – which is of course not itself a matter for stipulation – of instances of its [the principle's - Zs. N.] right hand side. […] All that is stipulated is the truth of a (universally quantified) biconditional". Hale (2001), 347, Hale and Wright (2002), 121. As far as I can see, however, this answer misses the point. The charge holds even if the existence of the intended abstract referents of the numerical terms of the identity statements on the left hand side of the instances of Frege's Abstraction Principle is stated conditionally, if and only if the equivalence statements on the right hand side of those instances are true. The concern is, obviously, why should a non-spatiotemporal entity exist if and only if a certain equivalence among certain spatiotemporal entities, such as lines and extensions of concepts, actually obtains? Without Frege's Abstraction Principle, and the referentialist construal of truth sanctioning the neo-Fregean inference from truth to the existence of subject matter, no such existence-claim could be derived.

[35] The separation of platonism about mathematical objects and properties from platonism about mathematical truth may also help Hale and Wright understanding the sense in which some platonists have found neo-Fregean platonism "insufficiently robust" in the philosophy of mathematics. As they say, "[t]he other accusation, that abstractionist platonism falls so far short of the genuine article as to be unworthy of the name 'platonism' at all, is harder to come to grips with, partly because it is hard to find a clear, articulate and non-metaphorical account of what 'genuine' platonism is supposed to involve and partly because any of several distinct things may lie behind it". Hale and Wright (2002), 121. According to my understanding, the robust platonist's problem with the neo-Fregean account is

however: the truth conditions of our beliefs about abstract domains can acquire a realist construal in a naturalistic framework as well. A naturalist conception of mathematical truth would, indeed, explain one of the core intuitions behind the neo-Fregean account as well: it would illuminate why our knowledge of abstract entities (in both of the earlier contrasted senses of the term) is no more problematic from Benacerraf's perspective than our knowledge of some spatiotemporal entities, such as those referred to on the right hand side of the instances of Frege's Abstraction Principle.

Summing up, although Hale and Wright present their case as a platonist, no-contact theorist response to Benacerraf's challenge in the philosophy of mathematics, their neo-Fregean account does not really explain how we can develop concepts of non-spatiotemporal entities in the first place, and (on a charitable reading) it cannot be reconciled with a platonist (i.e. realist and referentialist) construal of mathematical truth either. The adoption of a non-referentialist, naturalist construal of truth in the semantics of discourses about non-spatiotemporal domains, moreover, fits well with the neo-Fregean intuition that our knowledge of abstract objects and properties is no more problematic than our knowledge of the obtaining of some spatiotemporal states of affairs, but only in so far as it leaves open the possibility of developing a (naturalistic) contact theorist account of the relevant types of knowledge.

Three other platonist accounts of mathematical knowledge, which seem to provide a no-contact theorist answer to Benacerraf's challenge, involve a straight commitment to the platonist construal of the truth conditions, as opposed to merely the intended referents, of mathematical beliefs. According to the first, advocated, among others, by Jerrold Katz and David Lewis, our knowledge of abstract mathematical facts requires no epistemic contact with the relevant obtaining truth conditions,

that it does not support the idea that there are abstract objects and properties in the world and that our knowledge about these entities involves knowledge of their actual existence as well.

because this obtaining is necessary in character.[36] The reason for which we need to contact the intended referents of our empirical knowledge claims is that these entities could have been different. If a certain state of affairs necessarily obtains, then there is no point to check whether it in fact does so.

One may wonder, however, why the necessary character of mathematical facts would guarantee that our actual mathematical claims tend to be true, rather than false. The alleged fact that mathematical objects exist and are necessarily as they are, in itself, does not seem to guarantee or explain the reliability of our actual mathematical belief formation. If there is no information-conveying link between the obtaining truth conditions of our mathematical beliefs, on the one hand, and our actual evidence for these beliefs, on the other, then it seems that we have still no reason to suppose that our mathematical claims correctly represent the facts of the intended abstract realm. Consequently, the appeal to the necessary character of mathematical and other abstract facts does not seem to resolve the epistemological problem with the platonist construal of the relevant truths either.[37]

According to the second account, formulated, among others, by Michael Resnik and Stewart Shapiro, mathematical knowledge is knowledge of certain structures, which can be exemplified by various systems of abstract or spatiotemporal objects, and we can acquire this knowledge by constructing consistent axiom systems, because such systems provide implicit definitions of the structures characterised.[38] As Shapiro puts it, in relation to arithmetic:

[36] Katz (1981), Katz (1995), Lewis (1986).
[37] Field (1989), 238, Balaguer (1998), 41-45.
[38] Resnik (1997), Shapiro (1997), Shapiro (2000). Note that *in re* versions of structuralism with a physicalist background ontology (i.e. versions maintaining that the intended subject matters of mathematics are structures to be exemplified by systems of spatiotemporal entities) could *prima facie* provide a referentialist response to Benacerraf's dilemma, since referentialism, according to these versions, would imply naturalism or physicalism, rather than platonism about mathematical truth. As it has been mentioned in chapter 3, however, the idea that

> The structuralist vigorously rejects any sort of ontological independence among the natural numbers. The essence of a natural number is its *relations* to other natural numbers. The subject-matter of arithmetic is a single abstract structure, the pattern common to any infinite collection of objects that has a successor relation, a unique initial object, and satisfies the induction principle.[39]

The core constituents of this structuralist response to Benacerraf's challenge correspond to those of Hale and Wright's neo-Fregean account: consistent implicit definitions can provide us with both concepts and (*a priori*) knowledge of abstract entities (in this case, structures of various systems of objects) that exist in the world independently of our actual thoughts and knowledge of them; and the development of such definitions is something that we can explain without invoking an epistemic contact between our minds and the postulated abstract (non-spatiotemporal) entities.

The main problem with these referentialist (*ante rem* or *in re*) versions of structuralism is that, similarly to Hale and Wright's neo-Fregean platonism, they cannot account for our knowledge of the existence of the intended abstract subject matters, and thus for our knowledge of mathematical truths either. Knowledge of consistent implicit definitions cannot amount to, or imply, any knowledge of the existence of what the concepts defined purport

the subject matter of pure mathematics (and other discourses about abstract domains) is spatiotemporal in character is incompatible with the intentionalist construal of reference briefly put forward in section 4 of chapter 1. Furthermore, these versions of structuralism have troubles with the explanation of the apriority and necessity of the relevant truths, and the infinity of the relevant intended domains. Other structuralists, such as Benacerraf (1965), Putnam (1967) and Hellman (1989), embrace realism about mathematical truth without presupposing the existence of (non-spatiotemporal) mathematical entities. These positions may be in line with the non-referentialist framework advocated here, but they must be adjusted by a suitably articulated account of the factual basis of mathematical truths in the actual world.
[39] Shapiro (2000), 258.

to stand for, whether the purported entities are particular objects with intrinsic properties, or merely positions in an abstract structure that lack any individuating properties beyond their stipulated relations to other positions in the structure. The apriorist epistemology underlying these accounts can work only if the (realistically understood) truth of the relevant beliefs does not require the existence of the intended abstract subject matters, or in other terms, if truth in the semantics of these discourses is construed along non-referentialist lines. Again, such a construal would support the suggested apriorist account of knowledge by allowing an epistemologically unproblematic naturalist account of the truth conditions of the relevant beliefs. This would let us maintain the idea of objective knowledge about the intended non-spatiotemporal entities, and remain agnostic about the actual existence of these subject matters.[40]

Beyond failing as accounts of mathematical truth, structuralist forms of referentialism in philosophy of mathematics can be queried as construals of the subject matter of mathematics as well. If reference were to be understood in pure model-

[40] The "modal-structural interpretation" put forward in Hellman (1989) may be an example of this non-referentialist strategy in the philosophy of mathematics. According to this account, mathematics is about structures that could but actually may not be exemplified by particular systems of objects in the world. The standard objection to this account is that since mathematical truths are necessary in character, if some mathematical entities may exist, then they must exist, even in the actual world. Resnik (1992), 117, Shapiro (2000), 274, Isaacson (forthcoming). Note, however, that in a non-referentialist framework, the necessity of mathematical truths need not imply the necessity of the (metaphysically thick) obtaining of mathematical states of affairs. If the truth conditions of mathematics are not referential in character, then the idea that mathematics is about possibly existing entities that may not exist in the actual world becomes compatible with the view that mathematics consists of necessary truths. Of course, the objection remains valid, even in a non-referentialist framework, if our referential intentions in mathematics imply that the intended abstract subject matters in question must not exist unless they exist in the actual world. But why should one who has already abandoned referentialism about mathematical truth maintain anything about the mode of existence of mathematical objects and their properties? Such stipulations would be fully arbitrary in so far as they would no longer contribute to the explanation of any phenomenon in the actual world.

theoretic terms (i.e. in terms of "satisfaction"), then we would have, indeed, good reasons for construing mathematical objects and properties along the structuralist lines: a mathematical theory can be satisfied by various systems of objects (arithmetic, for instance, by both a system of Zermelo and a system of von Neumann ordinals), which could equally serve as the intended subject matter of the (syntactically understood) theory in question.[41] In the previous chapter, however, I argued that Putnam's permutation argument successfully demonstrates that our notion of reference or subject matter cannot be reduced to the pure model-theoretic notion of satisfaction. The fact that Peano's arithmetic is satisfied by various pluralities does not mean that our numerical concepts are about the structurally identified members of these pluralities, or about the positions themselves occupied by these members. Our numerical concepts are about numbers, which are meant to be different from both the set-theoretic entities invoked by Zermelo and von Neumann, and the structural positions that these entities occupy in the systems they are meant to constitute. The difference is manifest in our cognitive and linguistic practice as well: on the one hand, we do not think that the natural number 2 is identical with the set $\{\{0\}\}$ in the system of Zermelo ordinals or with the set $\{0,\{0\}\}$ in the system of von Neumann ordinals; on the other hand, we also deny that it is identical with a structural position that can be filled by various entities. In view of these cognitive and linguistic facts, it is hard to see what could justify the replacement of the traditional "object-platonist" construal of the subject matter of mathematics (and other discourses about non-spatiotemporal domains) with one of the structuralist alternatives.[42]

[41] The connection between this model-theoretic notion of reference and the appeal of a structuralist construal of mathematical entities is manifest in almost every representative paper of the structuralist tradition in the philosophy of mathematics.

[42] Structuralist construals of intended referents seem to be more adequate when we turn to our concepts of roles within a game or a system of institutions. The subject matter of our concepts of the white king's bishop or the President of the United States, for instance, is meant to be identical, in each context of application,

The third no-contact theorist account of mathematical knowledge that clearly endorses platonism about mathematical truth relies on a specific conception of the platonic realm. Mark Balaguer has baptised this view *plenitudinous* or *full-blooded platonism* (*FBP* for short), and it consists in the idea that all logically possible mathematical objects exist.[43] To see how this version of platonism is meant to supply an answer to Benacerraf's epistemological challenge, let me recall Balaguer's main objection to the "implicit definitionist" accounts of mathematical knowledge and reference presented above. The essential problem, on Balaguer's view, with Katz's and Lewis's necessity-based or Resnik's and Shapiro's structuralist replies is that none of them can explain how we could know that our implicitly defined mathematical concepts or internally consistent mathematical descriptions pick out an object in the alleged mathematical realm. As Balaguer puts it:

> Platonists can claim that the term '4' is just an abbreviation for the term 'successor of 3', but what anti-platonists will demand is an explanation of how we could know that there is an object in the mathematical realm that *answers* to this description. In other words: it's very easy to give definitions of mathematical singular terms like '4', but it's not so easy to see how we could know which terms and definitions actually refer to something.[44]

Or a few pages later:

either with a particular entity that occupies a certain role, respectively, in a certain game or in a certain system of institutions, or with the structural position itself occupied by the previous objects or individuals.

[43] Balaguer (1998). We must add that Balaguer's purpose in his book is not so much to defend platonism about mathematical truth and reference, but instead to show that in the philosophy of mathematics there are no conclusive arguments against either of the opposite doctrines of platonism and anti-platonism.

[44] Balaguer (1998), 42-43.

> Putting this response into the lingo that Resnik and Shapiro use, the problem is that prima facie, it seems that platonists cannot claim that we can acquire knowledge of abstract mathematical structures by merely formulating axiom systems that implicitly define such structures, because in making this claim, nothing is said about how we can know which of the various axiom systems that we might formulate actually pick out structures that exist in the mathematical realm.[45]

As Balaguer observes, however, if *FBP* is true, then there is a trivial platonist reply to this objection:

> For if all the mathematical objects that possibly *could* exist actually *do* exist, as FBP dictates, then all (consistent) mathematical descriptions and singular terms will refer, and any (consistent) representation of a mathematical object that someone could construct will be an *accurate* representation of an actually existing mathematical object.[46]

One may raise various objections to *FBP* and this reply to Benacerraf's challenge. Many of them are properly discussed and answered by Balaguer.[47] The most important problem, however, that he does not seem to appreciate is that what advocates of *FBP* can at most explain is how a reliable method of establishing logical relations (e.g. the consistency of alternative axiom systems) could also count as a reliable method of discovering platonic facts, *in case* the states of affairs in question obtain, indeed, in the real world. Still, they cannot tell us anything about how we could know that the latter condition holds after all. According to Balaguer, this is not a serious failure though, since:

[45] Balaguer (1998), 45.

[46] Balaguer (1998), 43.

[47] Balaguer (1998), ch. 3, esp. 58-69. For a sympathetic critical review of Balaguer's ideas, see Colyvan and Zalta (1999).

> [a]nti-platonists are not demanding here an account of how human beings could know that there exist any mathematical objects at all. That, I think, would be an illegitimate skeptical demand. […] All we can demand from platonists is an account of how human beings could know the *nature* of mathematical objects, *given* that such objects exist.[48]

I think that Balaguer is right when he distinguishes the task of showing how human beings could gain knowledge, or acquire reliable beliefs, about the nature of mathematical objects, given that such objects exist, from that of providing reasons for beliefs in the very existence of these entities. I also understand that there is a weak reading of Benacerraf's challenge, according to which the problem with platonism about truth is that it undermines the explanation of how our spatiotemporal belief-forming mechanisms *could* provide us with reliable information of platonic entities *if* the latter existed as platonists suppose. Finally, I grant that Balaguer's *FBP* amounts to an acceptable response to this reading of the challenge: it shows that under suitable circumstances our actual belief-forming mechanisms could provide us with largely true beliefs about platonic states of affairs, supposing that those indeed obtain. On the other hand, I believe that the intended reading of Benacerraf's challenge is stronger than the one suggested above. On this reading, what Benacerraf queries is the platonist's ability to explain how our spatiotemporal belief-forming mechanisms could provide us with reliable beliefs about what actually obtains in a platonic realm, including the issue of whether the truth conditions of *FBP* itself obtain.

In line with this reading, of course, I also query Balaguer's claim that the call for an account of how human beings could know that there are non-spatiotemporal entities in the metaphysically thick sense of the term is an illegitimate sceptical demand. In chapter 5, I argued that in the case of our beliefs

[48] Balaguer (1998), 43.

about the spatiotemporal world we can actually put forward an account of how human beings can know of the (thick) obtaining of the truth conditions or the existence of the subject matter of their beliefs. Further, earlier in this chapter, I argued also that in the case of our beliefs about fictive entities our referential intentions undermine the conceivability of such an account if we suppose that the truth conditions of these beliefs are referential in character, but the demand can be met if we adopt a non-referentialist, naturalistic construal of fictive truths. So, why should we abandon this demand in the philosophy of our discourses about abstract domains?

Balaguer is right: it would be a position on a par with radical scepticism if we did not allow the advocates of *FBP* (or other forms of platonism about mathematical truth) to assume at the beginning of their explanation that the relevant abstract entities exist. The charge, however, that anti-platonists raise against their opponents is not that they make initial assumptions of the intended non-spatiotemporal domains, but instead that the assumptions in question do not help us to understand why we should believe in their correctness at the end of the day.[49]

In response to this charge, advocates of *FBP* may invoke the explanatory virtues of platonism about the paradigms of *a priori* truth reviewed in section 1. In particular, they may argue that those virtues provide us with sufficient epistemic ground for believing in the correctness of a platonist's initial metaphysical assumptions. This reasoning, however, presupposes that the phenomena explained by reference to the suggested platonic facts cannot be properly explained otherwise (i.e. without invoking the existence of the relevant platonic objects and properties). In the following chapter, I shall show that this presupposition is false: by adopting a naturalistic, non-referentialist construal of the truth conditions of our beliefs about non-spatiotemporal domains, we

[49] As I argued in chapter 5, our standard assumption that there is an external, spatiotemporal world beyond the veil of our narrowly understood experience is one in terms of which we can explain why that experience can be reasonably regarded as a sign of the correctness of this assumption.

can explain every phenomenon that platonists intend to explain by invoking a non-spatiotemporal ontology. In view of these results, we can conclude that the metaphysical assumptions of an advocate of *FBP* cannot be based on the suggested explanatory considerations.

With this conclusion we have completed the survey of the most influential apriorist no-contact epistemologies in the philosophy of mathematics. As we saw, similarly to the empiricist form of this general strategy, none of these accounts can suitably explain how we could acquire knowledge of what obtains in a platonic realm. In a referentialist semantical framework, this failure undermines the accounts' adequacy as a theory of knowledge or reliable belief formation of non-spatiotemporal domains. The general problem with these attempts is that by denying the existence of an information-conveying link between our minds, on the one hand, and the (referentially construed) truth conditions of our beliefs about abstract domains, on the other, they deprive the latter conditions from any explanatory power *vis-à-vis* our mental life, so that nothing in our mind can be reasonably regarded as a distinctive evidence of the obtaining of the relevant truths.[50]

In view of this result, we shall conclude that Benacerraf's epistemological challenge cannot be properly answered by reference to a no-contact epistemology either. Together with our previous verdict on platonist contact epistemologies, this conclusion implies that the advocates of a referentialist (i.e. platonist) version of realism about truth in the semantics of our paradigm *a priori* discourses cannot account for the possibility of knowledge or reliable belief formation of the intended abstract

[50] In the conclusion of his book Balaguer also recognises this point. Balaguer (1998), 157. Nevertheless, instead of adopting a substantive realist and non-referentialist construal of mathematical truth, he rather endorses the somewhat innovative position that "*we* could never settle the dispute between platonists and anti-platonists", since "there is *no fact of the matter* as to whether platonism or anti-platonism is true, that is, whether there exist any abstract objects". Balaguer (1998), 152.

domains (i.e. their conception does not meet the fourth adequacy condition set for a theory of *a priori* truth in chapter 2).

Summary

In this chapter, I examined the prospects of the most influential platonist replies to Benacerraf's updated and generalised dilemma presented in chapter 3. The common feature of these responses is that they attempt to solve Benacerraf's problem by querying one of the epistemological premises of his case.

In section 1, I provided a brief overview of the major explanatory virtues and vices of platonism about truth in the semantics of our discourses about abstract domains, and argued that in absence of a viable account of knowledge acquisition or reliable belief formation about platonic realms one may query the adequacy of the advanced platonist explanations of apriority, intersubjectivity and observable convergence as well, which considerably reduces the explanatory power of the platonist construal.

In section 2, I turned to the discussion of the major platonist epistemologies. Following Mark Balaguer's useful distinction and terminology, I divided these accounts into two mutually exclusive and jointly exhaustive categories, and called them contact theories and no-contact theories of knowledge of platonic domains, respectively. First, I examined the contact theorist responses to Benacerraf's epistemological challenge, and argued that they are equally inadequate, since the explanation they provide of our knowledge or reliable belief formation about platonic domains is *ad hoc* and exotic in character. On the one hand, the adoption of such an epistemology would discourage any further inquiry into the nature of this type of knowledge acquisition, while on the other, it would open the door for parallel stipulations in the case of any other knowledge claims, no matter how the beliefs in question were causally produced in the subjects' minds. After this, I turned to the second group of platonist epistemologies, which query that our knowledge of abstract domains requires some information-conveying contact between our minds and the obtaining intended abstract truth conditions. My primary objection to these theories was that by denying the existence of the above contact, they deprive the suggested truth conditions from their explanatory significance *vis-à-vis* the actual constatations of human minds, so that nothing that we are aware of remains there to be reasonably regarded as a distinctive evidence of the obtaining of the relevant abstract truths.

It is important to emphasise that we found nothing objectionable in the platonist construal of the subject matter of our paradigm *a priori* beliefs about abstract domains. As we observed, our referential intentions in pure logic and mathematics allow for such a construal. According to these intentions, numbers and propositions are non-spatiotemporal entities, which may exist in the actual world, and if they exist there, then they exist in a platonic realm. What most platonists do not seem to realise is that platonism about subject matter does not imply platonism about truth. In particular, they fail to see that the conditions whose obtaining we take to be necessary and sufficient for certain logical and mathematical truths, and whose obtaining in the world we are supposed to know when we possess logical and mathematical knowledge, are not necessarily those that these beliefs purport to be about. Once we abandon the received referentialist construal of truth in the semantics of discourses about abstract domains, our platonism about the intended non-spatiotemporal referents will no longer stand in the way of a naturalist construal of the relevant truths, and a corresponding causal, contact theorist account of *a priori* knowledge or belief formation.

In chapter 3, I argued that there are four *prima facie* admissible theoretical options for a referentialist to escape Benacerraf's original or modified and generalised challenge in the philosophy of those discourses, in which we are supposed to acquire knowledge or reliable beliefs about causally inert domains. In the last three chapters, I examined these options, and showed that none of them can save the adequacy of standard

referentialism as a universal conception of truth. The two semantical responses (deflationism and anti-realism about truth) prove to be inadequate, since they cannot account for the objectivity of our knowledge claims, while the two epistemological responses (non-causal contact theories and no-contact theories of knowledge of realistically construed abstract domains) fail, because the account they provide is either *ad hoc* and exotic (non-causal contact theories) or insufficient (no-contact theories). In view of these results, we can conclude that Benacerraf's updated and generalised dilemma cannot be answered in a referentialist semantical framework.

With the fall of the above referentialist responses, there remains only one theoretical option for us to answer Benacerraf's dilemma: adopting a non-referentialist construal of truth in the semantics of our discourses about causally inert domains. In a non-referentialist framework, we may endorse a naturalistic conception of truth conditions while maintaining the received non-revisionist (i.e. platonist) construal of the intended subject matters. The resulting conception would be realist about truth, which means that it could explain the objectivity of this semantic property within these problematic discourses as well. On the other hand, by the naturalistic construal of the relevant truth conditions it would also support a causal, contact theorist account of how we can acquire knowledge or reliable beliefs about causally inert domains. Putting it briefly, a suitable non-referentialist response to Benacerraf's dilemma seems to satisfy those adequacy conditions that its referentialist alternatives failed to satisfy. The main question, of course, is whether the advocates of this response can also account for the other major *explananda* advanced in chapter 2. My central claim is that the answer to this question is positive. And it is this that I intend to show in the last chapter of this work.