

Forthcoming in: in Piotr Stalmaszczyk (ed.) *The Cambridge Handbook of the Philosophy of Language*, Cambridge University Press.

Truth and Theories of Truth

Panu Raatikainen

1. Introduction

The concept of truth and competing philosophical theories on what truth amounts to have an important place in contemporary philosophy. The aim of this chapter is to give a synopsis of different theories of truth and the particular philosophical issues related to the concept of truth. The literature on this topic is vast, and we must necessarily be rather selective and very brief about complex questions of interpretation of various philosophers. The focus of the chapter is mainly on selected systematic issues and the most influential and well-established philosophical theories and key concepts.¹

The idea that truth is some sort of correspondence with reality has been very common in the history of philosophy, though differing views have not been absent either. For a long time, however, there was neither much explicit debate nor more detailed accounts of the nature of truth, only some scattered remarks in the works of various classic philosophers. Many seem to have taken the correspondence idea as a given.

More well-developed contemporary theories of truth began to emerge at the beginning of the twentieth century, when Moore and Russell defended the correspondence view against (what they interpreted to be) the competing views on truth of some of their contemporaries, such as Joachim's coherence theory and James' pragmatist theory. Especially as reactions to the alleged problems of the correspondence theory, different alternatives to it have gained some popularity. In fact, there is quite a lot of overlap between the pragmatist, epistemic, and coherence theories of truth, and the views of many relevant philosophers include aspects from more than one such theory. In what follows, however, these theories will be discussed separately and systematically.

2. Truth-bearers

There are many different kinds of things to which truth is ordinarily attributed: beliefs, judgments, (declarative) sentences, utterances, propositions, etc. These are now commonly called 'truth-bearers'. However, some of these may inherit their truth from more fundamental truth-bearers. For example, perhaps an utterance is true only by virtue of expressing a true belief or true proposition. There have been different claims about what things are primary truth-bearers.

One natural candidate is the declarative *sentence*. Apparently, the sentence 'grass is green' is true, but the sentence 'snow is red' is false. What about sentences such as 'I am Swedish' or 'It is snowing here now'? They are not as such true or false; their truth-value depends on the

¹ Much more information on the considerably wider variety of philosophical and technical issues related to truth and theories of truth can be found in the sources mentioned in the Further Reading at the end of this chapter.

context of the utterance, e.g. on the speaker, time, and place. Consequently, it has been sometimes suggested that it is rather particular *utterances* in specific contexts, not sentences as types, that are primary truth-bearers. One problem with this view is that there exist (in some sense of ‘exist’) complex sentences which are never uttered but seem nevertheless to be true or false.

Others argue that the primary truth-bearer is a more abstract entity, a *proposition* or a ‘thought’ (in Frege’s sense). Propositions are variously characterised – in addition of being (allegedly) primary truth-bearers – as the meanings of sentences, ‘what is said’ or conveyed, and the contents of beliefs (and other attitudes). However, it is controversial whether such entities exist at all, and what sort of entities more specifically they are supposed to be. There are several quite different rival accounts of propositions on offer. All of these views have their own troubles. There is also the question of whether propositions should have and can have some sort of finer structure.²

The question of primary truth-bearers is a difficult philosophical problem. In what follows, we shall talk quite freely of sentences, beliefs, etc. as being true, whatever is natural in the context, without taking any stance on the debate.

3. Classical Correspondence Theory

The correspondence theory of truth is the view that the truth of a truth-bearer lies in its agreement, in some sense, with reality – or some portion of reality.³ For a long time, this was at least implicitly the default view in philosophy. The idea can already be found in the works of Plato and Aristotle. Brief expressions of it can also be found in the works of many key modern philosophers, such as Descartes, Locke, Spinoza, Leibniz, Hume, and Kant. However, the correspondence of a judgment was often taken by the modern classics to hold not with a fact or a state of affairs, but with the *object* of the judgment.

The now classical formulations in terms of *facts* were given by Russell: ‘Thus a belief is true when there is a corresponding fact, and is false when there is no corresponding fact’ (Russell 1912); and by Moore: ‘To say that this belief is true is to say that there is in the universe a fact to which it corresponds; and to say that it is false is to say that there is not in the universe any fact to which it corresponds’ (Moore 1953).

As such, the correspondence theory seems to be committed to the existence of a special category of entities, namely *facts*. However, it has remained controversial what kind of entities they are exactly supposed to be and whether it is at all plausible to postulate their existence. Moreover, given the complexity of some truth-bearers, the correspondence theory apparently needs to postulate some odd kinds of facts, such as negative or disjunctive facts. Many have taken this to be a fatal problem with the correspondence theory.

However, there is also another standard formulation of the correspondence theory that does not postulate facts in terms of *states of affairs*. According to this formulation, a truth-bearer is true if it corresponds to some state of affairs that obtains (see e.g. Austin 1950). This formulation, for its part, presupposes possible states of affairs that do not obtain to which false truth-bearers correspond – some may also find this suspect.

The correspondence theory has been vigorously criticised for other reasons too. There are complaints that the nature of the alleged correspondence relation is unbearably murky. Furthermore, it is popular to suggest that it is simply nonsensical to suppose that we could

² See Harris (this volume), Soames (this volume), Hanks (this volume), and Szabó & Thomason (2019, ch. 5.3.) for more about propositions.

³ David (2016) offers a good, thorough survey of both classical and contemporary correspondence theories.

somehow step outside our conceptual scheme or language and compare our beliefs or sentences directly with unconceptualised reality. Many have also worried that the correspondence theory opens a gap between reality and our knowledge of it and leads to scepticism (see David 2004, 2016 for further discussion). Further developments in correspondence theory are discussed in Section 9, below.

4. Coherence Theory of Truth

According to the coherence theory, the truth of a truth-bearer amounts not to correspondence with external reality, but to its coherence with a larger system of other truth-bearers.⁴ Often the focus has been on the truth of beliefs. This view has been attributed to, among others, Joachim (1906), Blanshard (1941), and Hempel (1935).

A standard critique of the coherence theory was famously presented by Russell (1906–1907): he proposed that it is likely possible to construct a coherent collection of sentences that would also cohere, for example, with the statement ‘Bishop Stubbs was hanged for murder’, which is however blatantly false. This is indeed a devastating objection for the view that truth consists of the mere internal coherence of a set of sentences or propositions in the abstract.

However, it is controversial whether any of the paradigmatic coherence theorists has really held this vulnerable view. It has been objected that in reality, the specific set with which the coherence is required consists of beliefs that are actually held. While this may rule out patent falsehoods, such as Russell’s statement about Bishop Stubbs, it is not clear that it really gets the coherence theory off the hook. First, the totality of beliefs held at any given time tends to contain beliefs that are, by any reasonable standard, false. A larger system of possible beliefs that coheres with them then necessarily also contains false beliefs. Second, even if one managed to avoid any false beliefs, there are typically indefinitely many mutually incompatible comprehensive sets of possible beliefs that are all equally coherent with the given finite set of actual beliefs. Obviously, all the beliefs in all these sets cannot logically be true, or otherwise contradictory beliefs would be true.

Defenders of the coherence theory have also tried to avoid such problems in the following ways: First, one might want to make ‘coherence’ something stronger than just logical consistency. For example, it has been suggested that ‘coherence’ consists of a particular belief’s being *entailed* by the other beliefs of the system. However, it is not clear that such a more demanding notion of coherence can be defined without an already presupposed notion of truth. Moreover, it is not obvious that even with such a strengthening, there could not be several mutually incompatible, equally coherent comprehensive sets of beliefs. There is still no well-developed account of such a stronger notion of ‘coherence’ to hand.

Second, it has been sometimes added that, in order to be true, beliefs must cohere with some specific, privileged set of beliefs – typically, beliefs about sense experiences. Then again, there are, in all reason, many mutually exclusive comprehensive sets of beliefs that cohere with all true beliefs about sense experiences, and clearly it would be contradictory to take all beliefs contained in all of them to be true.

Third, one might want to qualify the subject of the belief or the nature of the believer(s). After all, it is a fact of life that people and communities have conflicting beliefs about various matters; accordingly, the comprehensive coherent systems are also necessarily different for them. There is again a threat of unbearable relativism and bare inconsistency. Consequently, it has sometimes been suggested that those beliefs whose coherence is at stake are beliefs of an idealised subject or omniscient being, or that the focus is on the coherence of the set of beliefs

⁴ Walker (1989) provides a standard, book-length discussion of the coherence theory.

held at the ideal limit of inquiry. Such qualified, more defensible versions of the coherence theory come close to epistemic theories of truth, which will be discussed shortly.

5. Pragmatist Theories of Truth

The pragmatist tradition in philosophy was founded by Charles S. Peirce, but the pragmatist theory of truth is frequently attributed specifically to William James (1907, 1909). According to popular interpretation, James' view is that truth is what works, is useful, is expedient, or 'pays off' – i.e. something is true if and only if believing in it is useful.

The bedrock of pragmatist philosophy is its general view of concepts, which is encapsulated in 'the pragmatic maxim.' The maxim was first proposed by Peirce, and it contends that concepts are constituted by their connections to human practices and experiences: 'Consider what effects, that might conceivably have practical bearings, we conceive the object of our conception to have. Then, our conception of these effects is the whole of our conception of the object.' (Peirce 1878) The maxim was put forward as a method of clarifying concepts. James specifically believed that as a consequence, many major metaphysical debates would turn out to be idle, as it would make no practical difference which alternative one takes. James' understanding of the concept of truth is seemingly the result of this general maxim.

Upon closer scrutiny, it is unclear whether even James unequivocally identified truth with usefulness (see e.g. Putnam 1987), and other pragmatist philosophers such as Peirce or Dewey certainly did not. There are at least hints of the correspondence idea in the classic pragmatists, including James. There are also glimpses of the idea of coherence and various epistemic views on truth in James' writings. However, it cannot be denied that in some passages James does suggest something akin to what is conventionally taken to be 'the pragmatist theory of truth'. With such provisos in place, let us nevertheless consider the theory systematically. Some people have later found such a straightforward view attractive.

To begin with, it has been common to protest that the pragmatist theory conflates the definition and a criterion of truth. However, for pragmatists, this objection misses the point. From the perspective of the pragmatic maxim, they are not distinct. Of course, a critic could in turn challenge the maxim in general. The standard criticism of the simple pragmatist picture of truth dates back to Moore (1908). It contends that, although believing truths may indeed be often useful, its utility is neither sufficient nor necessary for a belief to be true. On the one hand, sometimes the truth may be so paralysing that it is harmful to dwell on it, and it is rather useful *not* to believe it. For example, confidence based on *false* beliefs about oneself may sometimes be useful. On the other hand, it is common to possess trivia that is completely useless, and wasting too much time and energy on acquiring such truths may even be harmful.

The pragmatist theory of truth does not at least amount to radical subjectivism: a person can be mistaken as to whether a belief is useful or not; hence a belief is not true simply because the subject believes it is. Nevertheless, as Moore also pointed out, intolerable relativism lurks behind the pragmatist view. Utility is often relative to an individual, a culture, and the circumstances. It may be useful for person *A* to believe that *p*, and useful for person *B* to believe not-*p*. Even for a single individual, several incompatible hypotheses may be equally useful and more useful than abstaining from judgment; therefore all would be 'true'. However, that is clearly unacceptable.

James attempted to respond to such concerns and insisted later that by 'expedient', he meant expedient 'in the long run and on the whole, of course' (James 1909). It is not clear, though, that this is really consistent with what he had previously said about truth. Be that as it may, we shall turn to more idealised ideas in the following section.

6. Epistemic Theories of Truth

Epistemic theories of truth contend that truth is constitutively connected to epistemic concepts such as verification, warrant, or justification: such epistemic concepts are taken as more primary, and truth is to be defined in terms of some such notion. The motivation has often been the desire to avoid scepticism. The correspondence view allows the possibility that even opinion in idealised epistemic circumstances – at the end of future inquiry or at the ideal limit – might be false. Epistemic theories emphatically deny this. They guarantee the knowability of truth by *defining* truth as knowability of some kind.

Sometimes James, for example, seems to say that a belief is true only inasmuch it has been *actually* verified. This is a truly radical view, as it makes truth drastically temporal: a belief becomes true only if and when it is verified. This differs from our ordinary understanding, according to which the belief must have then been true all along. Few have found such an extreme view appealing. Dummett (1978) has also argued that it is impossible to ground any reasonable logic in such a fluctuating conception of truth; a more stable notion seems necessary.

On other occasions, however, James took the more permissive view that a belief is true if it would be *possible* to verify it. More recently, Dummett (1978, 1993) has influentially advocated a similar view. Although his focus has often been on mathematical truth, and he has identified truth in that context with provability, Dummett has also proposed that truth in general should be equated with verifiability in principle. This notion has played a central role in Dummett's semantic anti-realism and verificationist semantics (see Harris (this volume), Dummett 1978, 1993).

A related, more idealised epistemic view of truth was held already by Peirce, who argued truth is 'the opinion which is fated to be ultimately agreed to by all who investigate' (Peirce 1878). Putnam, in his 'internal realism' period (circa 1976–1990), advocated a view not unlike the aforementioned one: 'Truth is an idealization of rational acceptability. We speak as if there were such things as epistemically ideal conditions, and we call a statement "true" if it would be justified under such conditions.' (Putnam 1981)

More sophisticated epistemic theories of truth have been quite popular, but they also face problems. To begin with, *verification* – whether actual or idealised and 'in principle' – seems unfit for the role of the primary concept here: many truths, e.g. more theoretical truths in science, are not even verifiable in principle; at best, they can only achieve confirmation of some degree through observation. Confirmation of a scientific claim by observation is furthermore holistic and depends on the other claims of the theory and various auxiliary assumptions, which may vary.

The basic problem with all epistemic views of truth is that *evidence* can be misleading. For example, all the available evidence may point to the conclusion that a person who is in fact innocent is guilty. As a response, it may be added that one must focus on definitive, final, non-defeasible evidence. However, can the latter be defined without the concept of truth? Moreover, there are plausibly many specific truths about the distant past for which all evidence has perished. No amount of future scientific inquiry, however long it is pursued and the evidence gathered, would be sufficient for the view to determinately converge to such a truth.

Although reality is for Peirce independent of any individual subject's beliefs and hence objective, it may not be entirely mind-independent: he sometimes apparently identifies reality with whatever is postulated by the final opinion. On the other hand, the reliability of the final consensus (our approaching it) and the scientific method is explained, according to Peirce, by the fact that reality controls our observations. This seems to entail that something from the future has an impact on our current experiences – but that is absurd.

Peirce and many of his followers advocate fallibilism as a plausible midway between desperate scepticism and the unrealistic requirement of infallible and certain knowledge. A part of this picture is the idea that the scientific method is 'self-corrective' and can approach and

approximate the truth. However, if the truth is in turn defined as the limit of scientific inquiry, the said central fallibilist idea becomes trivial and empty (see Laudan 1973).

Justification, in general, is quite relative. Therefore Putnam, for example, added the requirement of *epistemically ideal conditions*. But is it possible to explicate the notion of such ideal justification without already presupposing the notion of truth? Candidates that first come to mind, such as *a reliable indicator of truth*, suggest this may not be possible. If so, the whole account is circular.

All the same, it seems plausible that in an ideal situation, the subject must possess all evidence relevant to the statement, for otherwise some additional evidence might defeat the statement. On closer scrutiny though, the whole idea of *all relevant evidence* is baffling. First, as Alston has pointed out, all beliefs somehow relevant to x likely constitute a contradictory set of beliefs, and restricting attention only to all *true* beliefs would already presuppose the concept of truth (Alston 1996). Second, as Wright (1992) has noted in this context, it is common knowledge that justification is vastly holistic – i.e. whether some evidence is relevant for something depends essentially on background assumptions, and it is thus relative. In sum, it remains unclear whether the concept of all evidence relevant to some statement is even a meaningful concept.

As problems such as those mentioned above have accumulated, epistemic theories have become less popular.

7. Formal Approaches to Truth

In the early twentieth century, philosophers were commonly sceptical about the very meaningfulness of the concept of truth. Awareness was growing of various logical paradoxes and anomalies related to truth, such as the liar paradox.⁵ In addition, more philosophical reasons were being given for this aversion.⁶

The atmosphere changed dramatically with Alfred Tarski's pioneering logico-mathematical contribution to the topic (Tarski 1935; see Patterson 2012; Hodges 2018). He aimed to provide a rigorous definition of 'true' without presupposing any semantical concepts.⁷ To begin with, Tarski emphasised the importance of clearly distinguishing what he called 'the object language' from 'the metalanguage'. The former is the language which is 'talked about', and the latter is the language *in which* we talk about the first language. Tarski argued that the object language and the metalanguage cannot coincide, or one ends up with paradoxes. Accordingly, it is not possible to give a fully general rigorous definition of truth, but only to define a truth predicate ' x is true-in- L ' for a particular object language L at one time. It is not permissible, in Tarski's setting, to iterate the truth predicate; it is admissible to apply it only to sentences of the object language L that do not contain the truth predicate for L . With the Tarskian restrictions in place, paradoxical sentences cannot even be formulated: they are not well-formed sentences at any level.

What Tarski did was to show that, assuming (i) the syntax of the object language L is specified exactly enough, (ii) the reference of its primitive expressions (names and predicates)

⁵ The liar paradox refers broadly to the variety of claims that somehow attribute falsity to themselves. They all tend to entail logical contradiction. The simplest version is 'This sentence is false,' but a myriad of less direct ways of referring to the claim itself might be used, e.g.:

'The last sentence of the 5th footnote [of "Truth and theories of truth"] is false.'

⁶ For some history of the philosophical views on truth before Tarski, see e.g. Niiniluoto (1999), Sluga (1999), and Woleński & Simons (1989).

⁷ By 'semantical concepts', Tarski meant, roughly, concepts that involve relations between linguistic expressions and the objects about which these expressions speak.

is given (in the metalanguage), and (iii) the metalanguage has a modest amount of set-theoretic power, one can explicitly define a truth predicate for L . Furthermore, what can be explicitly defined can be eliminated. It follows that the defined concept cannot give rise to any inconsistencies (i.e. paradoxes). This gave new respectability to the concept of truth.

Tarski put forward his famous ‘Convention T ’ as the criterion of material adequacy for any proposed definition of a truth predicate. He required that any adequate definition should entail as its consequence all instances of the schema:

$$(T) \quad X \text{ is true} \leftrightarrow P,$$

where ‘ X ’ is the name of a sentence S of the object language in the metalanguage, and ‘ P ’ is a translation of that sentence S into the metalanguage. For example, if the object language is a fragment of German and the metalanguage is English, the following is an instance of the schema (T):

$$\text{‘Schnee ist weiss’ is true-in-German} \leftrightarrow \text{snow is white.}$$

Alternatively, if the metalanguage ML is a direct extension of the object language L , and both are well-defined fragments of English, one could have, e.g.:

$$\text{‘Grass is green’ is true-in-}L \leftrightarrow \text{grass is green.}$$

Such instances of the schema (T) are nowadays commonly called T -sentences.⁸

Tarski’s theory does not impose any epistemic constraints on truth and is in line with the correspondence theory: a truth may well be unknowable in it. However, the question of whether Tarski’s account as such amounts to a full-blown substantial correspondence theory of truth in contradistinction to deflationist views on truth (see below) is more complicated. On the one hand, Tarski emphatically did not propose that mere T -sentences would constitute a sufficient theory of truth, as is typically held by contemporary deflationists. Several philosophers have interpreted Tarski’s full account of truth as a version of the correspondence theory. On the other hand, Field (1972) has emphasised that Tarski’s truth definition rests on a list-like definition of the *denotation* of primitive expressions; according to Tarski, a definition for names could look something like the following (where the object language L is a fragment of German):

$$\begin{aligned} \text{Denotes}_L(x, y) \leftrightarrow \\ & [(x = \text{‘Frankreich’} \wedge y = \text{France}) \vee \\ & (x = \text{‘Deutschland’} \wedge y = \text{Germany}) \vee \\ & \quad \vdots \\ & (x = \text{‘Köln’} \wedge y = \text{Cologne})]. \end{aligned}$$

As Devitt (2001) notes, ‘such list-like definitions are in no way explanatory, but are essentially deflationary and so could not yield anything substantial about reference.’ Consequently, Devitt maintains that Tarski’s definition of truth itself does not show us anything substantial about truth; rather, it has an overall somewhat deflationist flavour. However, if Tarski’s account were revised by dropping its list-like definitions and then supplemented by a substantial theory of reference, we would have a genuine correspondence theory of truth.

⁸ Tarski’s theory and T -sentences play a fundamental role in Davidson’s truth-conditional semantic programme. However, whereas Tarski presupposed meanings (translation) and aimed to explicate the notion of truth, Davidson reversed the direction; by presupposing the concept of truth, he hoped for an explication of sentence-meanings with the help of truth-conditions provided by T -sentences. See Harris (this volume) and Davidson (1984).

Successful as it has been, Tarski's approach places severe restrictions on the truth predicate; it must be relativised to a particular object language L_1 , and if one wants to talk about the metalanguage L_2 and the truth of its sentences, a further metalanguage L_3 and a distinct truth predicate is needed, and so on. The Tarskian theory thus leads to a (sort of type-theoretic) hierarchy of different languages and truth predicates. It is common to say that such truth predicates are 'typed'.

However, as Kripke (1975) has pointed out, there are many natural uses of 'true' in cases where it does not make sense to assign a determinate Tarskian language or type to it (e.g. 'indirect ascriptions'; see below). In Kripke's own new technically sophisticated approach, the general truth predicate is not typed and its application is not restricted; in particular, the truth predicate can be iterated. It is then possible to formulate paradoxical sentences, but they are now 'ungrounded' and lack a truth-value. Martin and Woodruff (1975) independently presented a very similar approach. Since then, the variety of different formal theories of truth on the market has been an embarrassment of riches.

These more general theories come with a price though. Whatever its limits, Tarski's approach does not deviate at all from classical logic. In theories like Kripke's, so-called 'inner logic' and 'outer logic' must be distinguished. Roughly, the former refers to the logic that sentences 'inside' the truth predicate obey, whereas the latter refers to the logic outside the applications of the truth predicate. In recent non-typed theories, these two logics often come apart; for example, both $S \vee \neg S$ and $\neg \text{True}('S \vee \neg S')$ may be provable for some S .

As Leitgeb (2007) has concisely outlined, there are a number of desiderata that we would naturally expect a formal theory of truth to satisfy. They include, e.g.:

- (i) The truth predicate should not be subject to any type restrictions;
- (ii) T -equivalences should be derivable unrestrictedly;
- (iii) The outer logic and the inner logic should coincide;
- (iv) The outer logic should be classical.

However, it is a hard logico-mathematical fact that no possible theory can satisfy all these desiderata. It is necessary to compromise on some of them. In particular, if one sticks to (i), then both (ii) and (iv) cannot be simultaneously satisfied.⁹ Tarski, of course, gave up (i). It is far from clear what choice would be ideal, and more philosophical work is required here.

8. Deflationist and Minimalist Theories

The neo-classical theories of truth discussed above all share the assumption that whatever its essence, truth is a genuine, substantial property. Deflationist and minimalist theories deny this: they hold that truth has no 'nature' to be revealed, nor is truth a substantial property that could do some explanatory work.¹⁰

The most straightforward theory in this family is the traditional *redundancy theory of truth*, often attributed to Ayer (1936). It contends that *all* talk of truth can be *always* eliminated: according to it, if P is a declarative sentence, to say that ' P ' is true is simply to say that P . Intuitive equivalences that resemble Tarski's T -sentences – let us call them T -equivalences – play an essential role here.¹¹

⁹ Some corners have been cut here; for the exact details, the reader is recommended to consult Leitgeb's (2007) original paper, which is very readable and illuminating.

¹⁰ Stoljar & Damnjanovic (2014) offer a good overview of deflationist views and the alleged challenges for them.

¹¹ However, Tarski's T -sentences only involve material equivalence, whereas deflationism typically holds that these equivalences hold in virtue of analyticity or, sometimes, even synonymy. Furthermore, some deflationists impose no Tarskian-type restrictions for the truth predicate and allow its iteration.

However, it is nowadays well known that so-called ‘*indirect ascriptions of truth*’ and *generalisations involving truth* cannot be that easily eliminated; this was indeed the stumbling block of the traditional redundancy theory. By indirect ascriptions of truth, one means uses of truth such as ‘What Frank said yesterday is true’, ‘Fermat’s last theorem is true’, or ‘That is true’. By generalisations involving truth, uses such as the following are intended: ‘Every sentence that is logically entailed by a true sentence is true’, ‘All theorems of the theory *F* are true’, ‘Everything the Pope says is true’, etc. In general, the sentences to which truth is ascribed are not explicitly given in such cases. In such contexts, the truth predicate cannot just be eliminated with the help of *T*-equivalences. Present-day deflationists and minimalists grant this and add that these and related (what they call ‘logical’) contexts are the only contexts where the notion of truth is needed; it plays no further explanatory role. Nevertheless, the standard formulations of deflationism contend that *T*-equivalences exhaust the concept of truth and constitute our theory of truth.

Field (1986)¹² has suggested that at least most of the uses of ‘true’ that have proved difficult to paraphrase away are uses where ‘true’ is serving as a surrogate for *infinite conjunction* or *infinite disjunction* (in which ‘true’ would not occur). Consider, for example: ‘There are true sentences which no one will ever have grounds to accept’. This use of ‘true’ can, according to Field, be naturally understood as a device of infinite disjunction:

*P*₁, but no one will ever have grounds to accept it; or
*P*₂, but no one will ever have grounds to accept it; or
 :

Inspired by Quine (1970),¹³ Field (1986) called this notion ‘a notion of disquotational truth’. At the time, Field suggested that even someone who accepts a correspondence notion of truth needs this notion. He then defined ‘a deflationary notion of truth’ as the view that proposes the latter serves no useful purpose at all while at the same time preserving a use for the word ‘true’, contrary to the radical redundancy view. A few years later, Horwich published a book-length defence of deflationism; he called his own particular brand of deflationism ‘minimalism’ (Horwich 1990). The main difference from the above definition is that he took propositions as the primary bearers of truth, but besides that, he generally agreed with Field’s characterisation. According to Horwich, ‘the truth predicate exists solely for the sake of a certain logical need’, namely, the problem of having a single, finite proposition that has the intuitive logical power of an infinite conjunction. The (unproblematic) instances of the schema

The proposition *that p* is true $\leftrightarrow p$

form what Horwich calls ‘the minimal theory of truth’ (the epithet ‘non-problematic’ is intended to rule out possible paradoxical instances). Horwich contends that all facts involving truth can be explained on the basis of the minimal theory. By 1994, Field himself had become convinced of the correctness of deflationism as he had earlier defined it (see Field 1994). Since then, the view has had numerous sympathisers.

In accordance with Eklund (2019), the key claims of contemporary deflationism can be summarised as follows:

- (a) Truth is not a genuine, substantial, or explanatory property.

¹² At the time, Field had not yet been converted to deflationism; nevertheless, Field (1986) offers one of the most useful discussions of deflationism.

¹³ Although Quine’s (1970) remarks on truth considerably inspired contemporary deflationism, it is problematic to count Quine himself unqualifiedly as a deflationist.

- (b) The truth predicate exists solely for the sake of a certain logical need.
- (c) The concept of truth is exhausted by the *T*-equivalence schema, and all there is to know about truth follows from the instances of the schema.

Strictly speaking, we must distinguish the official deflationist theory of truth consisting of (unproblematic) *T*-equivalences from the general deflationist background view typically comprising claims (a) – (c).

Many have found deflationism appealing. However, it is not without its own problems (cf. Stoljar & Damnjanovic 2014). To begin with, it is a powerful *intuition* that truth amounts to correspondence with reality. Deflationism must at least accommodate this intuition. It is not completely clear how it can do this, though various responses have been attempted. Further, it is common to think that truth is a norm of belief and assertion – that they aim at truth. It has been frequently suggested that deflationism cannot account for this *normativity* of truth. It is nevertheless quite difficult to make this objection more exact. *Paradoxes* are something that any theory of truth faces. It has been argued, however, that deflationism has special difficulties in solving the paradoxes of truth. It has also been contended that it is particularly challenging for deflationism to handle ‘*truth-value gaps*’, i.e. sentences that are neither true nor false – if such are allowed.

At least when focusing on sentences, truth depends on *meanings*. But how can a deflationist explicate them? Due to the threat of circularity, the popular truth-referential accounts of meaning seem to be out of question. Thus, an alternative is required. At present, the proposed use-theoretic theories of the leading deflationists (see Field 1994, Horwich 1998) are highly programmatic at best. Resorting to propositions may not harmonise well with the anti-metaphysical spirit of deflationism, and it anyway brings with it the difficult question of the more specific nature of propositions. In any case, their explanation cannot apparently utilise truth-referential concepts.

As noted, one central thesis of contemporary deflationism is that the sole purpose of truth predicates is to be able to express certain *generalisations* (and indirect ascriptions). Yet now that they can be expressed, how can they be known to be true? It would be obviously implausible to require that one’s theory of truth (in deflationism: *T*-equivalences) entail *all* generalisations involving truth, for some such generalisations are not even true: consider ‘Everything that Trump says is true’. However, in the case of some very general facts involving truth, seemingly independent of any contingent facts, such as:

For all sentences *S*: $True('S') \rightarrow \neg True(' \neg S')$,

it is difficult to see how they could be achieved if one’s theory of truth does not entail them. The standard deflationist theory that consists of *T*-equivalences is impotent to deliver them; it only provides instances of the generalisation.¹⁴ This is now standardly called ‘the generalisation problem’. It is a pressing challenge for deflationism at least as it is generally understood.¹⁵

There is also a lack of clarity concerning what exactly deflationism is a theory of. Is it a theory about a property, or a theory about a linguistic predicate or a concept that corresponds to a property? Deflationism contends that truth is not a substantial or explanatory *property*. However, it proceeds to a claim about the function of the *predicate* that it serves solely logical purposes. Furthermore, what do *T*-equivalences constitute a theory of? (See Eklund 2019; cf. Devitt 2001.)

It seems that deflationist theories are now in flux: the standard versions based on *T*-equivalences face troubles. Still, many find deflationism’s general minimalist and anti-

¹⁴ This was first pointed out in the present context by Gupta (1993). The observation itself goes back to Tarski (1935). By contrast, Tarski’s full theory does entail such generalisations.

¹⁵ For some further moves, see Raatikainen (2005, 2006) and Armour-Garb (2010).

metaphysical attitude attractive. Perhaps novel theories will emerge that are still deflationist in spirit and more defensible, but that remains to be seen.

9. Developments in the Correspondence Theory

In spite of all the opposition, the correspondence theory has not vanished; it is still perhaps the most popular view among philosophers. Let us take a brief look at how it has evolved since its classical versions.

It is helpful to distinguish between weak and strong correspondence theories (Woleński & Simons 1989), or, correspondence-as-congruence and correspondence-as-correlation (see Pitcher 1964; Kirkham 1992). According to the *weak correspondence theories* or *correspondence-as-correlation* views, every truth-bearer is correlated to a state of affairs, and if that state of affairs to which a given truth-bearer is correlated actually obtains, the truth-bearer is true; otherwise it is false.¹⁶ The *strong correspondence theories* or *correspondence-as-congruence* (or, sometimes, *correspondence-as-isomorphism*) views further require that there is a structural isomorphism between truth-bearers and the facts or states of affairs to which they correspond if true. A truth-bearer mirrors or pictures the state of affairs to which it is correlated. Nothing of the sort is assumed by the former, weaker idea of correspondence, which holds that a truth-bearer as a whole is correlated with a state of affairs as a whole. Weak correspondence involves only the idea that truth depends somehow on how things are in the world.

Wittgenstein (1921) and Russell (1918) advocated *logical atomism*. According to this idea, there are simple atomic or elementary truth-bearers whose truth amounts to correspondence (in the strong congruence sense; Wittgenstein called it ‘picturing’) with atomic facts. However, the truth-value of more complex truth-bearers is a matter of language, and it is determined by the truth-bearers’ logical structure and the truth-values of their atomic parts, along the lines of truth tables. Consequently, their logical atomism does not postulate – as earlier Russell and Moore held – any complex and bizarre (e.g. disjunctive and negative) facts, only atomic ones. As such, it avoids one powerful objection to correspondence theories.

The correspondence relation has been accused of being intolerably mysterious. One might counter that it is no more murky than semantic relations (e.g. think about something, refer to something) in general – which are very difficult to completely get rid of in philosophy. Whatever unanswered questions there are here, the problem is not specific to the theory of truth. As we have noted, it is possible to take Tarski’s formal theory¹⁷ as a model and reduce the truth of even atomic sentences to reference of their sub-sentential parts, names, and predicates: what object a name denotes or refers to, and to which objects a predicate applies, as Devitt (2001) and the early work of Field (1972), for instance, suggest. For example, the sentence ‘*n* is *P*’ is true if ‘*n*’ refers to *o* and ‘*P*’ applies to *o*. This strategy avoids presupposing even atomic facts. Such a Tarski-inspired theory only assumes correspondence in the weak correlation sense.

In the philosophy of language, there are motives to develop a plausible theory of reference for names and predicates independent of any worries about theories of truth. Although the existing theories are still sketchy, some progress has been made, and such theories can play some explanatory role (see Martí (this volume)). It is then just a short step to formulate a theory of truth for sentences along broadly Tarskian lines that is based on such referential relations. One limitation of such an approach is that it is unclear how exactly it could be extended to cover

¹⁶ For simplicity, the formulation in terms of states of affairs is used here; however, the general idea in no way requires this particular formulation.

¹⁷ The formal theories of Tarski and Kripke (and others) do not usually differ much here except when it comes to iterating the truth predicate.

all sorts of sentences involving broadly modal contexts. Therefore, as such, it is at best a simple model for a more general correspondence theory.

On the whole, the trend among the advocates of the correspondence theory has been away from the stronger congruence idea towards the weaker correlation view. The theory of Austin (1950) is an already classic example. Instead of naïve ideas of ‘picturing’ or ‘mirroring’ reality, weaker and more flexible conceptions of correspondence have been suggested. Goldman (1986), for example, has proposed the metaphor of ‘*fittingness*’ – that language or thought corresponds to reality like clothes fit a body; indefinitely many different kinds of trousers may fit one’s body, but many others just do not fit. Recycling Ramsey’s old metaphor, Kitcher (2001) has compared scientific theories to *maps* such as the underground map of London. Depending on its purpose, such a map is highly selective and involves conventional elements such as colours. It would be absurd to think that it comprehensively mirrors London and is the whole truth about it. The idea of a uniquely correct map makes no sense. Nevertheless, a map can be more or less accurate for its purpose. Kitcher suggests that our theories correspond (or not) to the world in a somewhat analogous way.

As for the worries about stepping outside our language or conceptual scheme, an advocate of the correspondence theory could counter that her theory requires no such thing: the objection conflates the definition of the general property *true* (or the meaning of ‘true’) and our ways of coming to know which particular truth-bearers are actually true.

However, even if the correspondence theory could be formulated rigorously and defended against all the critiques, the question of whether there is any actual demand for such a more substantial notion of truth – in contrast to the deflationist notion – still remains. For some attempts to argue that there is, see the work of Devitt (2001) and Kitcher (2002). This is, though, still very much an ongoing debate.

Further Reading

For more about theories of truth, see e.g. Kirkham 1992, Schmitt 1995, Engel 2002, Künne 2003, David 2004, and Glanzberg 2018a, 2018b. The relevant chapters of Haack 1978 are still helpful. Many central sources are collected in Blackburn & Simmons 1999, Lynch 2001, and Schmitt 2004. For more about formal approaches to truth, see e.g. Leitgeb 2007; Horsten & Halbach 2014; Beall et al. 2018; Burgess & Burgess 2011.

References

- Alston, William P. (1996). *A Realist Conception of Truth*, Ithaca and London: Cornell University Press.
- Armour-Garb, Bradley (2010). ‘Horwichian Minimalism and the Generalization Problem,’ *Analysis* 70, 693–703.
- Austin, J. L. (1950). ‘Truth’, *Proceedings of the Aristotelian Society*, Supplementary Volume 24: 111–128; reprinted in J.L. Austin, *Philosophical Papers*, 3rd ed., Oxford: Oxford University Press 1979, 117–33.
- Ayer, A. J. (1936), *Language, Truth, and Logic*, 2nd ed., 1946, New York: Dover.
- Beall, Jc, Michael Glanzberg & David Ripley (2018). *Formal Theories of Truth*, Oxford: Oxford University Press.
- Blackburn, Simon & Keith Simmons, eds. (1999). *Truth*, Oxford: Oxford University Press.
- Blanshard, Brand (1941), *The Nature of Thought*, vol. 2, New York: The Macmillan Company.
- Burgess, Alexis G. & John P. Burgess (2011). *Truth*. Princeton: Princeton University Press.
- David, Marian (2004). ‘Theories of Truth.’ In I. Niiniluoto, M. Sintonen & J. Woleński, eds., *Handbook of Epistemology*, Dordrecht: Kluwer Academic Publishers, 331–414.

- David, Marian (2016). 'The Correspondence Theory of Truth.' *The Stanford Encyclopedia of Philosophy* (Fall 2016 Edition), Edward N. Zalta, ed., URL = <https://plato.stanford.edu/archives/fall2016/entries/truth-correspondence/>.
- Davidson, D. 1984: *Inquiries into Truth and Interpretation*. Oxford: Oxford University Press.
- Devitt, Michael (2001). 'The Metaphysics of Truth.' In Lynch, ed. (2001), 579–611.
- Dummett, Michael (1978). *Truth and Other Enigmas*, London: Duckworth.
- Dummett, Michael (1993). *The Seas of Language*, Oxford: Clarendon Press.
- Eklund, Matti (2019). 'What is Deflationism about Truth?,' *Synthese*. (forthcoming)
- Engel, Pascal (2002). *Truth*, Chesham: Acumen.
- Field, Hartry (1972). 'Tarski's Theory of Truth,' *The Journal of Philosophy* 69, 347–75.
- Field, Hartry (1986). 'The Deflationary Conception of Truth.' In G. MacDonald & C. Wright, eds., *Fact, Science and Morality*, Oxford: Blackwell.
- Field, Hartry (1994). 'Deflationist Views of Meaning and Content,' *Mind* 103, 249–84.
- Glanzberg, Michael (2018a). 'Truth.' *The Stanford Encyclopedia of Philosophy* (Fall 2018 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/fall2018/entries/truth/>.
- Glanzberg, Michael, ed. (2018b). *The Oxford Handbook of Truth*, Oxford: Oxford University Press.
- Goldman, Alvin (1986). *Epistemology and Cognition*, Cambridge, MA: Harvard University Press.
- Gupta, Anil (1993). 'A Critique of Deflationism,' *Philosophical Topics* 21, 57–81.
- Haack, Susan (1978). *Philosophy of Logics*. Cambridge: Cambridge University Press.
- Hempel, Carl (1935). 'On the Logical Positivists' Theory of Truth,' *Analysis* 2, 49–59.
- Hodges, Wilfrid (2018). 'Tarski's Truth Definitions.' *The Stanford Encyclopedia of Philosophy* (Fall 2018 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/fall2018/entries/tarski-truth/>.
- Horsten, Leon and Volker Halbach (2014). 'Truth and Paradox.' In L. Horsten & R. Pettigrew, eds., *The Bloomsbury Companion to Philosophical Logic*, London and New York: Bloomsbury, 351–82.
- Horwich, Paul (1990). *Truth*, Oxford: Basil Blackwell. (2nd edition, Oxford: Clarendon Press 1998).
- Horwich, Paul (1998). *Meaning*, Oxford: Clarendon.
- James, William (1907). *Pragmatism: A New Name for Some Old Ways of Thinking*, New York and London: Longmans, Green & Co.
- James, William (1909). *The Meaning of Truth: A Sequel to "Pragmatism"*, New York and London: Longmans, Green & Co.
- Joachim, Harold H. (1906) *The Nature of Truth*, Oxford: Clarendon Press.
- Kirkham, Richard L. (1992). *Theories of Truth: A Critical Introduction*, Cambridge, MA: MIT Press.
- Kitcher, Philip (2001). *Science, Truth, and Democracy*, Oxford: Oxford University Press.
- Kitcher, Philip (2002). 'On the Explanatory Role of Correspondence Truth,' *Philosophy and Phenomenological Research* 64, 346-364.
- Kripke, Saul (1975). 'Outline of a Theory of Truth,' *The Journal of Philosophy* 72, 690–716.
- Künne, Wolfgang (2003). *Conceptions of Truth*, Oxford: Clarendon Press.
- Laudan, Larry (1973). 'Charles Sanders Peirce and the Trivialization of the Self-corrective Thesis', in R. Giere & R. Westfall (eds.), *Foundations of Scientific Method in the 19th Century*. Bloomington: Indiana Univ. Press, 275–306.
- Leitgeb, Hannes (2007). 'What Theories of Truth Should Be Like (But Cannot Be),' *Philosophy Compass* 2 (2), 276–290.
- Lynch, Michael P., ed. (2001). *The Nature of Truth: Classical and Contemporary Perspectives*, Cambridge, MA: MIT Press.
- Martin, R.L. and P.W. Woodruff (1975). 'On Representing "True-in-L" in L,' *Philosophia* 5, 213–17.
- Moore, G. E. (1908). 'Professor James' "Pragmatism".' *Proceedings of the Aristotelian Society* 8(1), 33–77.
- Moore, G. E. (1953). *Some Main Problems of Philosophy* (lectures given in 1910–11). London: George Allen & Unwin.
- Niiniluoto, Ilkka (1999). 'Theories of Truth: Vienna, Berlin, and Warsaw.' In J. Woleński and E. Köhler (eds.), *Alfred Tarski and the Vienna Circle*, Dordrecht: Kluwer, 17–26.
- Patterson, Douglas (2012). *Alfred Tarski: Philosophy of Language and Logic*, New York: Palgrave Macmillan.

- Peirce, C. S. (1878). 'How to Make Our Ideas Clear.' *Popular Science Monthly* 12 (January 1878), 286–302.
- Pitcher, George (1964). 'Introduction.' In G. Pitcher, ed., *Truth*, Englewood Cliffs: Prentice-Hall, 1–15.
- Putnam, Hilary (1981). *Reason, Truth and History*, Cambridge: Cambridge University Press.
- Putnam, Hilary (1997). 'James's Theory of Truth.' In Ruth Anna Putnam, ed., *The Cambridge Companion to William James*, Cambridge: Cambridge University Press, 166–185.
- Quine, W.V. (1970). *Philosophy of Logic*, Englewood Cliffs, Prentice Hall.
- Raatikainen, Panu (2005). 'On Horwich's Way Out,' *Analysis* 65, 343–346.
- Raatikainen, Panu (2006). 'Problems of Deflationism.' In T. Aho & A. Pietarinen, eds., *Truth and Games in Logic and Language*, Acta Philosophica Fennica Vol. 78, Helsinki: Societas Philosophica Fennica, 175–85.
- Russell, Bertrand (1906–07). 'On the Nature of Truth,' *Proceedings of the Aristotelian Society* 7, 28–49.
- Russell, Bertrand (1912). *The Problems of Philosophy*, London: Williams and Norgate.
- Russell, Bertrand (1918). 'The Philosophy of Logical Atomism.' In *Logic and Knowledge: Essays 1901–1950*, R. C. Marsh, ed., London: George Allen and Unwin, 1956, 177–281.
- Schmitt, Frederick F. (1995). *Truth: A Primer*, Boulder, Colorado: Westview Press.
- Schmitt, Frederick F. (ed.) 2004. *Theories of Truth*, Oxford: Blackwell.
- Sluga, Hans (1999). 'Truth before Tarski.' In J. Woleński & E. Köhler, eds., *Alfred Tarski and the Vienna Circle*, Dordrecht: Kluwer, 27–41.
- Stoljar, Daniel & Damnjanovic, Nic (2014). 'The Deflationary Theory of Truth', *The Stanford Encyclopedia of Philosophy* (Fall 2014 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/fall2014/entries/truth-deflationary/>.
- Szabó, Zoltan Gendler & Richmond H. Thomason (2019). *Philosophy of Language*. Cambridge: Cambridge University Press.
- Tarski, Alfred (1935). 'The Concept of Truth in Formalized Languages.' In A. Tarski: *Logic, Semantics, Metamathematics* (2nd edition) J. Corcoran ed., Hackett, Indianapolis, 1983, 152–278.
- Walker, Ralph C. S., (1989). *The Coherence Theory of Truth*, London: Routledge.
- Wittgenstein, Ludwig (1921). *Tractatus Logico-Philosophicus*. English translation (by D. F. Pears & B. F. McGuinness), London & Henley: Routledge & Kegan Paul 1961.
- Woleński, Jan & Peter Simons (1989). 'De Veritate: Austro-Polish Contributions to the Theory of Truth from Brentano to Tarski.' In K. Szaniawski, ed., *The Vienna Circle and the Lvov–Warsaw School*, Dordrecht: Kluwer, 391–442.
- Wright, Crispin (1992). *Truth and Objectivity*, Cambridge, MA: Harvard University Press.