

Truth

Theories of truth investigate truth as a property of our thoughts and speech. We attribute truth and falsity to a wide variety of so-called *truth-bearers*: linguistic items (sentences, utterances, statements, and assertions), abstract items (propositions), and mental items (judgments and beliefs). What is the property we are attributing when we call a truth-bearer true? The question is crucial because of truth's involvement in central philosophical claims: it is often said, for example, that truth is the aim of science, or that the meaning of a sentence is given by the conditions under which it is true, or that logical validity is the preservation of truth, or that ethical statements are neither true nor false. A proper understanding of truth promises to illuminate fundamental issues in metaphysics, the philosophy of language, logic and ethics.

The two traditional theories of truth are the correspondence theory and the coherence theory. Further theories of truth have emerged since the last part of the nineteenth century, most notably the pragmatic theory, the identity theory, and the semantic theory. There has also been a reaction against the idea that truth has a substantive nature to uncover, which has led to markedly increased support for so-called 'deflationary' theories of truth.

A very different motivation for theorizing about truth is the challenge posed by the semantic paradoxes, especially the Liar paradox. Theories of truth prompted by the Liar tend to be concerned less with the nature of truth, and more with the logic and semantics of the predicate 'true'. There has been surprisingly little contact between these two groups of theories (though see Beall and Armour-Garb (eds.) 2005).

The correspondence theory of truth

According to the correspondence theory, truth consists in correspondence to the facts. A truth-bearer (say, the proposition that snow is white) is true if and only if it corresponds to a fact (the fact that snow is white). Broadly speaking, truth is a relational property between truth-bearers on the one side and the world on the other.

There is the suggestion of the correspondence account in Plato's *Sophist* (263b), where in Theaetetus' presence, the Stranger contrasts the true statement 'Theaetetus sits' with the false statement 'Theaetetus flies': "the true one states about you the things that are as they are ... [w]hereas the false statement states about you things different from the things that are". Aristotle, in *Categories*, writes "The fact of the being of a man carries with it the truth of the proposition that he is ... the truth or falsity of the proposition depends on the fact of the man's being or not being" (14b14-22; see also 4b8). The correspondence idea may also be present in Aristotle's famous definition of truth (in *Metaphysics* Γ, 1011b25): "to say of what is that it is, and of what is not that it is not, is true". Echoes of the Platonic-Aristotelian account are present in the Stoics and medieval philosophers (for example, Aquinas, Ockham and Buridan), and many modern philosophers from Descartes onwards endorse the correspondence idea, though with little or no discussion.

A classic statement of the correspondence theory is given by G.E. Moore: to say of a given belief that it is true "is to say that there is in the Universe *a* fact to which it corresponds" (1910-11, p.302). Moore takes it that we are all perfectly familiar with the relation of correspondence:

“That there *is* such a relation, seems to me clear; all that is new about my definitions is that they concentrate attention upon just *that* relation, and make it the essential point in the definitions of truth and falsehood” (p.304)

Moore’s remarks bring out both a strength and a weakness of the correspondence theory. The correspondence theory is the most natural account of truth – it seems that no-one need deny that a true belief corresponds to how things are. But this raises the suspicion that the correspondence theory is platitudinous – to say that a truth bearer corresponds to the facts is just an elaborate way of saying that it is true. There is no distinctive *theory* of truth unless more can be said about the correspondence relation. And Moore admits that he can offer no analysis of it; the best he can do, he says, is to “define it in the sense of pointing out what relation it is, by simply pointing out that it is *the* relation which does hold between this belief, if true, and this fact, and does not hold between this belief and any other fact” (p.301).

Bertrand Russell (1906-1907, 1912) attempted to shed light on the correspondence relation by arguing for a structural isomorphism or congruence between beliefs and facts. Beliefs and facts are structured complexes, and when a belief-complex is suitably congruent with a fact-complex, the belief is true. Consider Othello’s belief that Desdemona loves Cassio. According to Russell, believing is a four-place relation; in the present case, it is the cement that unites Othello, Desdemona, the loving relation, and Cassio into one complex whole. The last three items are what Russell calls the *objects* in the belief, and these objects are ordered in a certain way by the believing relation (Othello believes that Desdemona loves Cassio, not that Cassio loves Desdemona). Now consider another complex unity, *Desdemona’s love for Cassio*, composed of the objects

in Othello's belief. Here the loving relation is the cement that binds together Desdemona and Cassio in the same order that they have in Othello's belief. If this complex unity exists, then it "is called the fact corresponding to the belief. Thus a belief is true when there is a corresponding fact, and is false when there is no corresponding fact." (p.129)

### Objections to the correspondence theory

It is central to Russell's elucidation that there is a structural congruence between the content of a true belief and the corresponding fact – for example, between the proposition expressed by the sentence 'Desdemona loves Cassio' and the fact that Desdemona loves Cassio. But sentences and the propositions they express come in a variety of logical structures – negations, conditionals, universal generalizations, and so on. Are there, then, 'funny facts': negative facts, hypothetical facts, universal facts, and other logically complex facts? It might seem that the real world – the world of dated, particular events and things in specific spatial and temporal orderings – just does not seem able to contain anything of this kind of complexity: negative or universal or hypothetical situations, for example. We seem to be presented with a dilemma: either facts are too 'linguistic', too closely tied to the logical structures of our language, or facts are worldly items that are not structurally congruent with the propositions we express.

Russell (1918) and Ludwig Wittgenstein (1921) went on to develop their philosophy of logical atomism, according to which there are no logically complex facts, only atomic facts. True propositions that are logically simple or atomic correspond to atomic facts, but logically complex true propositions no longer correspond to logically complex facts. Rather, complex propositions are recursively broken down into the simple

propositions which compose them, and the truth of complex propositions is ultimately explained via the atomic facts to which true atomic propositions correspond. Difficulties remain, however: certain complex propositions, for example ‘because’ statements and subjunctives, are resistant to a recursive breakdown into simple components; and we can still ask whether universal facts are required for true universal generalizations, and negative facts for true negations. Despite these well-known problems, versions of logical atomism are not without their supporters – see for example, David Armstrong (1997). In a different vein, J.L Austin avoids ‘funny facts’ by denying that correspondence is a matter of structural congruence: “There is no need whatsoever for the words used in making a true statement to ‘mirror’ in any way, however indirect, any feature whatsoever of the situation or event” (p.155) – even a single word or simple phrase can correspond to a complex situation. Rather, correspondence is a correlation that is determined by our linguistic conventions: it is “*absolutely and purely conventional*” (p.154).

A far-reaching and influential family of objections to the correspondence theory takes issue with a certain distinction of standpoints that the theory seems to imply. There is the standpoint we occupy when we judge, say, that there are cows in the garden, and then there is the standpoint we occupy when we determine whether our judgment is true. When we occupy this latter standpoint, the correspondence theory seems to require us to judge whether our judgment is appropriately related by correspondence to the facts. Frege (1918) objects that there really is no further standpoint to take up, and no further judgment to make – rather we should simply verify whether there are cows in the garden. This line of thought led Frege to the conclusion that truth is undefinable; it also tends towards deflationism, since it may seem that truth drops out of the picture. According to

another line of objection, it is an illusion that we can have access to an unvarnished realm of facts with which to compare our judgment. Our knowledge of the world is mediated by our descriptions, interpretations, and judgments; we cannot step outside our own system of beliefs and compare those beliefs with ‘bare reality’. Since the correspondence theory says that truth consists in correspondence to the facts, and those facts are inaccessible to us, we can never know that a judgment of ours is true, and we are led to skepticism. Those who endorse this line of criticism typically associate the correspondence theory with metaphysical realism, and advocate instead some form of anti-realism and an ‘epistemic’ account of truth, say in terms of verification (like the logical positivists) or assertibility (like John Dewey, and later, Michael Dummett).

### The coherence theory of truth

If we cannot judge a belief against the facts, perhaps we should judge it against our other beliefs: does it ‘hang together’ with the rest of our beliefs? The coherence theorist says that the truth of a belief consists in its coherence with other beliefs. Given some favored coherent set of beliefs, the truth of any of its members consists in its membership in that set – in this way the skeptic is disarmed, since truth no longer requires access to an independent realm of facts. Versions of the coherence theory have been attributed to Spinoza, Kant, Fichte and Hegel (cf. Walker 1989), and the theory was championed by idealists, including H.H. Joachim and Brand Blanshard, at the end of the nineteenth century and the beginning of the twentieth. Joachim (1906) rejects Descartes’ idea that we can know truths individually: “The ideal of knowledge for me is a system, not of *truths*, but of truth” (p.48); knowledge of an individual truth “is the smallest and

most abstracted fragment of knowledge, a mere mutilated shred torn from the living whole in which alone it possessed its significance” (*ibid.*). So Joachim advocates a thoroughgoing *holistic* view of knowledge, and of truth: “Truth in its essential nature is that systematic coherence which is the character of a significant whole” (p.50). The coherence theory was subsequently adopted by some logical positivists, notably Otto Neurath (1932-1933), who, like Joachim, endorsed a holistic view of knowledge and truth, and combined it with the positivists’ verificationist doctrine that no sense can be attached to a reality that goes beyond what can be verified or falsified by the empirical methods of science.

There are attractive features of the coherence theory: in favor of holism, we can say that statements like “The Enlightenment brought about the French Revolution” and “Neutrinos lack mass” cannot be understood in isolation from a good deal of history and science; and we do often *test* the truth of a statement against a large body of background statements. But the coherence theory is a theory of the *nature* of truth, not a theory of how we test for truth, and as such it has been the target of a number of objections. Russell (1906), Moritz Schlick (1934) and others have argued that an arbitrary set of propositions, say those of a fairy story or a good novel, would count as a set of *truths* as long as the propositions cohere with one another – where coherence is taken in the sense of consistency or compatibility. An appeal to comprehensiveness seems not to help the coherence theorist here: given a coherent set of propositions however large, there will always be equally large coherent sets incompatible with it (and with each other). And placing restrictions on membership in the favored set – for example, admitting only our actual beliefs, or ideal beliefs held at the end of inquiry – seems to tie truth less to

coherence and more to the successful tracking of the facts. A further objection derives from Russell: suppose we have a large, coherent set of propositions about, say, the nineteenth century, and suppose that we can coherently add the proposition that Bishop Stubbs wore episcopal gaiters. According to the coherence theory this proposition is true, in virtue of its membership in a coherent set. If we protest that we cannot be committed to its truth because we don't know whether it's true or false, then we are using true and false in a way that the coherence theorist doesn't recognize. The difficulty is compounded if we now run the argument with the proposition that Bishop Stubbs *didn't* wear episcopal gaiters. Further discussion of the coherence theory can be found in Putnam (1981), Blackburn (1984), Davidson (1984), and Walker (1989).

### The pragmatic theory of truth

The pragmatic theory of truth is associated primarily with the American pragmatists Charles Peirce and William James, and their influence can still be felt in the work of, for example, Richard Rorty and Robert Brandom. According to Peirce (1878), we are to understand any idea or object through its practical effects:

“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.” (p.31)

Peirce applies his rule to the idea of *reality*: the practical effect that real things have on us “is to cause belief” (p.36), and so the question is how to distinguish true belief from false belief. Peirce's answer is that the true beliefs are the ones to which we will all agree, and



only the methods of science can realize the hope of reaching this consensus. Peirce writes:

“This great hope is embodied in the conception of truth and reality. The opinion which is fated to be ultimately agreed to by all who investigate, is what we mean by the truth, and the object represented in this opinion is the real. That is the way I would explain reality.” (p.38)

This is not the independently existing reality associated with the correspondence theory: for Peirce, what is special about science is its ability to settle opinion, and reality is whatever settled opinion says it is.

James applies Peirce’s rule directly to truth. The practical effects of true beliefs are successful actions, beneficial dealings with the world; truths are “invaluable instruments of action” (James 1907, p.97), truths “pay” (p.104). And so, in accordance with Peirce’s rule, truth *is* what is useful, what ‘works’. James places less emphasis than Peirce on consensus and scientific method (indeed, Peirce renamed his theory ‘pragmatism’ to distance it from James’ version). James applies his theory to individuals’ beliefs as well as collective beliefs, and religious and metaphysical beliefs as well as empirical ones. (For example: “On pragmatist principles, if the hypothesis of God works satisfactorily in the widest sense of the word, it is true” (p.143)).

It is standardly objected that we can have beneficial false beliefs and detrimental true beliefs. My false belief that I play the violin beautifully may in fact improve my performance; my true belief that I don’t may worsen it. James has the resources for a response: while “the true is only the expedient in our way of thinking”, truth is the expedient in a strong sense: “expedient in the long run and on the whole of course”

(p.106). We have to take the long view: I may perform well this time, but overall I will better served by an accurate assessment of my talents. The long view must be taken not only of individuals' beliefs, but of whole theories – Ptolemaic astronomy was expedient for centuries (p.107).

“The ‘absolutely’ true, meaning what no farther experience will ever alter, is that ideal vanishing-point towards which we imagine that all our temporary truths will some day converge. ... Meanwhile we have to live today by what truth we can get today, and be ready tomorrow to call it falsehood”. (pp.106-7)

Along with a controversial commitment to relativism, James presents here a holistic theme which may suit his pragmatism: it is perhaps more plausible that the truth of an entire system of belief, as opposed the truth of our beliefs taken individually, is a matter of its working for us. Taken this way, pragmatism may be seen as a version of the coherence theory. Still, a basic objection remains: it is plausible that a body of truths should be useful or coherent, but it does not follow that truth *consists in* utility or coherence - a correspondence theorist will say that truths are useful and mutually coherent just because they correspond to the world.

### The identity theory of truth

Despairing of the correspondence theory, Bradley wrote: “if we are to advance, we must accept once for all the identification of truth with reality” (1907, pp.35-6). Here Bradley seems to embrace the identity theory of truth: a truth does not correspond to a fact, but is *identical* to a fact (Bradley's view is discussed in Candlish 1995). Another influence is Frege's remark: “A fact is a true thought” (1918, p.101), though Frege

himself did not endorse the identity theory. Versions of the theory are defended by Jennifer Hornsby (1997) and Julian Dodd (2000). The theory may appear quite counterintuitive: if true mental items – true judgments or true beliefs – are facts, then it seems that the mind *contains* facts, that mind and world are literally the same. It may also be argued that the theory is unstable, collapsing into deflationism, or leading to the elimination of true judgments altogether – “straight to thought’s suicide”, as Bradley puts it (1893, p.150).

### The semantic theory of truth

The semantic theory of truth originates with the mathematician and logician Alfred Tarski (1930-1931, 1944). Tarski sought a *definition* of truth that was formally correct and met the following constraint: it must imply all sentences of the form exemplified by

“Aardvarks amble” is true if and only if aardvarks amble,

that is, all sentences of the form: “p” is true if and only if p. These so-called “T-sentences” are so basic to truth, Tarski thought, that they must follow logically from any adequate definition – in this way, he says, we do justice to Aristotle’s definition (see above). Indeed, Tarski regards each T-sentence as a “partial definition” of truth, and if we were dealing with a finite language (in the sense that it contains only finitely many sentences), we need only list all the associated T-sentences for a complete definition of truth for that language (see 1930-1, pp.251-3). But since Tarski was after a definition of truth for formal languages that were infinitary, such a list is not feasible. So instead Tarski provided a recursive definition - not of truth, though, but of the more basic notion

of *satisfaction*. In the simplest kind of case, satisfaction is a relation between an object and a predicate - for example, a London bus satisfies the predicate 'is red'. Satisfaction is defined recursively, first for predicates (of a given language) that exhibit no logical complexity, and second for those that do. Tarski then defined truth in terms of satisfaction. The result was a definition of truth for formal languages that was formally precise and implied the T-sentences.

It is remarkable that both correspondence theorists and deflationists have found Tarski's account congenial. Correspondence theorists are drawn to satisfaction as a word-world relation, and the possibility that the correspondence relation between a sentence and a fact can be broken down into relations between parts of sentences (predicates and names) and the things they refer to (e.g. Devitt 1991). This raises the hope that correspondence is no more mysterious than the semantic relations between predicates and names and their referents. Deflationists, in particular disquotationalists, are drawn to the idea that the T-sentences say all there is to say about truth, as we will see below. Tarski himself emphasized the neutrality of his theory:

“we may accept the semantic conception of truth without giving up any epistemological attitude we may have had; we may remain naïve realists, critical realists or idealists, empiricists or metaphysicians – whatever we were before.

The semantic conception is completely neutral toward all these issues.” (1944, p.140).

Tarski's aim was not to uncover the nature of truth, but rather to place the concept of truth beyond suspicion. On the one hand, he thought, truth is fundamental to science, logic and metamathematics; on the other hand, truth has an “evil reputation” because of

its involvement with the Liar paradox. Tarski's aim was to find a way of defining truth in terms that no-one could question:

“the definition of truth, or of any other semantic concept, will fulfil what we intuitively expect from every definition; that is, it will explain the meaning of the term being defined in terms whose meaning appears to be completely clear and unequivocal. And, moreover, we have then a kind of guarantee that the use of semantic concepts will not involve us in any contradictions.” (1944, p.127).

Anyone wishing to turn Tarski's definition into a fully general account of truth faces a number of obstacles: Tarski defines truth only for regimented, formal languages, not for natural languages like English; the definition is a definition of truth for a given language, not for truth *simpliciter*; and the definition, according to Field (1972), fails to explain truth since it merely reduces truth to further semantic notions that are not themselves adequately explained.

### Deflationary theories of truth

Deflationists say that ‘substantive’ theories of truth – like the correspondence and coherence theories – are radically misguided: there is no substantive property of truth to theorize about. According to Frank Ramsey (1927), truth is *redundant*: “it is evident that ‘It is true that Caesar was murdered’ means no more than that Caesar was murdered” (p.106). Truth is less easily eliminated from generalizations like ‘Everything Socrates says is true’, but Ramsey argues that it can be done (*ibid.*). The word ‘true’ disappears, and any reason to investigate the nature of truth disappears along with it. According to a more sophisticated version of the redundancy theory, *the prosentential theory of truth*

(Grover, Camp and Belnap 1975), the word ‘true’ is not even a genuine predicate, but a mere component of *prosentences*. If I say “That is true” in response to a claim of yours, I have produced not a sentence but a prosentence, referring back to your sentence just as the pronoun ‘he’ may refer back to the name ‘John’. We might think of “That is true” as hyphenated, with no more internal structure than the pronoun ‘he’. On the prosentential view, ‘true’ does not survive as a discrete property-denoting predicate. P. F. Strawson’s variant of the redundancy theory (1949) attributes to ‘true’ a performative role: we use ‘true’ not to pick out a property, but to perform speech-acts such as endorsing, agreeing and conceding.

Disquotationalists also ascribe to ‘true’ a role very different from that of ordinary predicates. According to the disquotational theory of truth – championed by W.V. Quine (1970) and further developed and defended by Hartry Field (see e.g. 1994) – to say that a sentence is true is really just an indirect way of saying the sentence itself. There really is no more to the truth of the sentence ‘Penguins waddle’ than is given by the Tarskian T-sentence

“Penguins waddle” is true if and only if penguins waddle,

and the totality of T-sentences tells the whole story about truth. This prompts the question: why not dispense with the truth predicate in favor of direct talk about the world? The disquotationalist will respond by pointing to generalizations such as “Every sentence of the form ‘p or not p’ is true” (see Quine 1970, pp.10-13). In such a case, we could dispense with the truth predicate only if we could produce an infinite conjunction of sentences of the form ‘p or not p’: “Aardvarks amble or aardvarks do not amble, and bison bathe or bison do not bathe, and ...”. But we cannot produce infinitely long

sentences. So to achieve the desired effect, we generalize over sentences, and then, via the truth-predicate, bring them back down to earth by disquoting them. The truth-predicate is a device for disquotation. Despite surface appearances, ‘true’ does not denote a property or relation – it is a logical device. So there is no property of truth to explore, and no work for truth to do beyond its logical role.

The disquotational theory takes the truth-bearers to be sentences, and this raises a concern about the scope of the theory (for further concerns, see David 1994). Suppose that on the authority of others I believe that Dmitri is always right, though I speak no Russian. I say, with apparent understanding, “What Dmitri says is true”. But according to disquotationalism, understanding what I have said is just a matter of understanding what Dmitri said; and since I cannot understand what Dmitri said, I cannot understand what I have said. Disquotationalists typically relativize their theory to the sentences of a given natural language such as English. And since an English speaker will not understand every sentence of English, some disquotationalists recognize the need to go further and restrict the theory to the sentences of a given speaker’s idiolect (those sentences that the speaker understands). This seems to lead us away from the commonsensical notion of truth – ordinarily, it seems, we can apply the notion of truth to foreign sentences, and to sentences of English that we do not yet understand. In short, the concept of truth seems not to depend on the sentences that a speaker happens to understand at a given time. The challenge to the disquotationalist (taken up by Field and others) is to ease the counterintuitive restrictions on disquotational truth in ways that do not compromise the theory.

These difficulties for disquotationalism might motivate a different choice of truth-bearer – propositions instead of sentences. Paul Horwich (1990, 1998) presents a *minimal theory of truth*, according to which a complete account of truth is given by the propositional analogues of Tarski's T-sentences:

The proposition that aardvarks amble is true if and only if aardvarks amble

The proposition that bison bathe is true if and only if bison bathe,

and so on, *ad infinitum*. Far from being restricted to speakers' idiolects, 'true' applies to *all* propositions, including those expressed by sentences we don't understand. But now there is a new set of concerns. First, since we do not understand every proposition, we will understand only a fraction of the axioms that compose the minimal theory - and so our grasp of truth must always remain partial. Second, since the minimal theory describes truth in a piecemeal way, for each proposition individually, it does not include any *generalizations* about truth. So it may be objected that the theory cannot explain generalizations such as 'Only propositions are true' - the theory doesn't tell us what isn't true, so it doesn't rule out, for example, the absurdity that the Moon is true. (For more on this objection, see Gupta 1993; Hill 2002 offers a version of minimalism that is responsive to it.) Third, consider the form shared by Horwich's axioms: the proposition that p is true if and only p. To obtain an axiom, we must be careful to replace each occurrence of 'p' by English tokens of the same sentence-type, with the same meaning. But now sentences appear to be back in the picture – together with the substantive semantic notion of meaning, which may not be as free of involvement with truth as minimalism requires.



This last remark relates to a general challenge faced by all forms of deflationism. Deflationists typically focus on uses of ‘true’ such as “‘Aardvarks amble’ is true”, or “Most of what Socrates says is true” – what we may call *first-order* uses, where ‘true’ applies to a particular truth-bearer or a set of truth-bearers. But ‘true’ is also used in other ways: consider, for example, the claim that the meaning of a sentence is given by its truth-conditions, or the claim that to assert is to put forward as true. These uses of ‘true’, call them *second-order*, purport to explain meaning and assertion. Unlike first-order uses, they do not apply to any particular truth-bearers, and so it is not easy to see how they might be treated as redundant and eliminable, or given a merely disquotational role. These second-order uses must be explained. Moreover, the deflationist must show that it is possible to explain meaning and assertion (and many other concepts apparently related to truth, such as validity, belief, verification, explanation, and practical success) in terms that assign to truth a limited logical role or no role at all.

### Theories of truth and the Liar

One version of the Liar paradox is generated by the self-referential sentence:

(1) (1) is false.

Suppose that (1) is true, then what it says is the case, and so (1) is false. Suppose on the other hand that (1) is false - then since that’s what (1) says it is, (1) is true. We reach a contradiction either way: we are landed in paradox.

Hierarchical theories of truth have perhaps been the orthodox response to the Liar. Let  $L_0$  be a fragment of English that does not contain the predicate ‘true’. Let ‘true-in- $L_0$ ’ be the truth predicate for  $L_0$ , holding of exactly the true sentences of  $L_0$ . If ‘true-in-

$L_0$ ' is itself a predicate of  $L_0$ , then we can construct the Liar paradox in  $L_0$  via the sentence "This sentence is not true-in- $L_0$ ". Accordingly, the predicate 'true-in- $L_0$ ' is confined to a richer *metalanguage* for the *object language*  $L_0$ . But on pain of the Liar, this metalanguage cannot contain *its* own truth predicate; for that a further metalanguage is needed. In this way, a hierarchy of languages is generated, each language beyond  $L_0$  containing the truth predicate for the preceding language. By a celebrated theorem of Tarski's (proved in Tarski (1933)), no classical *formal* language can contain its own truth predicate, and we are led to a hierarchy of formal languages. Some have carried over this result to natural languages as a way of dealing with the Liar, though Tarski does not endorse this move. Russell's hierarchical approach is embodied in his theory of types and orders (1908). It is often complained that hierarchical approaches force an unnatural regimentation on a natural language like English; Russell himself at one time called the approach "harsh and highly artificial".

Another kind of approach abandons classical semantics – usually it is the principle of bivalence ('Every sentence is true or false') that is rejected. If we can motivate the existence of truth-value gaps, then we can say that (1) is neither true nor false, and avoid the contradiction. Saul Kripke's influential theory of truth (1975) takes liar sentences to be 'gappy' because they are *ungrounded*: any attempt to evaluate a liar sentence leads only to sentences involving 'true' or 'false' – in the case of (1), we are repeatedly led back to (1) itself. Kripke constructs a language which, remarkably, contains its own truth and falsity predicates. It cannot, however, accommodate the predicates 'is false or gappy' or 'not true' – and so ultimately we cannot dispense with a hierarchy.

The revision theory of truth (Gupta and Belnap 1993) is formally a variant of Kripke's theory, but provides a distinctive way of explaining the meaning of truth. Truth is taken to be a circular concept, and the revision theory describes how its meaning is given by the Tarskian T-sentences via a dynamic process which, through systematic revisions, provides better and better approximations of the extension of 'true'.

Contextual theories of truth are motivated by so-called "strengthened" reasoning about the Liar. Start with a Liar sentence, say

(2) (2) is not true.

Reasoning in the usual way, we will find that (2) is pathological. But then we may infer

(3) (2) is not true.

Now (2) and (3) are composed of the very same words with the very same meanings, and yet one is pathological and the other is true. Contextual theorists claim that this change in truth status without a change in meaning is best explained by a contextual shift (compare 'I'm hungry' said before dinner and 'I'm hungry' said after dessert). Most contextual theories are hierarchical (e.g. Burge 1979 and Barwise and Etchmندی 1987), though Simmons (1993) develops a suggestion of Gödel's, according to which an unstratified concept of truth applies everywhere except for certain singularities.

Any purported solution to the Liar faces the so-called Revenge Liar – a version of the Liar couched in the very terms of the solution. Truth-value gap approaches must deal with the liar sentence 'This sentence is false or gappy', hierarchical approaches with 'This sentence is not true at any level', and contextual theories with 'This sentence is not true in any context'. With no agreed-upon solution in sight, and with the constant threat of revenge liars, some have concluded that we must cut the Gordian knot and embrace the

contradictions associated with the Liar. According to *dialetheists* such as Graham Priest (1987) there are sentences that are *both* true and false, and among them are the liar sentences. (For critical discussions of dialetheism, see Priest *et. al.* (eds.) 2004). Besides meeting the obvious charge of counterintuitiveness, dialetheists must underwrite their theory with a plausible *paraconsistent* logic (a logic which challenges the principle that everything follows from a contradiction), and ensure that dialetheism is not itself vulnerable to a revenge liar.

Bibliography

- Alston, William P. *A Realist Conception of Truth*. Ithaca and London: Cornell University Press, 1996.
- Aristotle. *Categories and Metaphysics*, in *The Basic Works of Aristotle*, edited and with an introduction by Richard McKeon. New York: Random House, 1941.
- Armstrong, David. *A World of States of Affairs*. Cambridge: Cambridge University Press, 1997.
- Austin, J. L. "Truth." *Proceedings of the Aristotelian Society* suppl.vol.24 (1950): 111-129; repr. in Blackburn and Simmons (eds.) 1999: 149-161 (page references are to this reprint).
- Barwise, Jon, and Etchemendy, John. *The Liar*. Oxford: Oxford University Press, 1987.
- Beall, J.C., and Armour-Garb, Brad (eds.). *Deflationism and Paradox*. Oxford: Oxford University Press, 2005.
- Blackburn, Simon and Simmons, Keith (eds.). *Truth*, in the series Oxford Readings in Philosophy. Oxford: Oxford University Press, 1999.
- Blackburn, Simon. *Spreading the Word*. Oxford: Clarendon Press, 1984.
- Blanshard, Brand. *The Nature of Thought*. London: George Allen and Unwin, 1939.
- Bradley, F.H. *Appearance and Reality*. Oxford: Oxford University Press, 1893.
- Bradley, F.H. "On Truth and Copying". *Mind* 16 (1907):165-180; repr. in *Essays on Truth and Reality*, Oxford: Clarendon Press, 1914: 107-26; also repr. in Blackburn and Simmons (eds.) 1999: 31-45 (the page reference is to this reprint).
- Brandom, Robert. *Making it Explicit*. Cambridge, Mass.: Harvard University Press, 1994.

- Burge, Tyler. "Semantical Paradox". *Journal of Philosophy* 76 (1979): 169-198;  
repr. in Martin (ed.) 1984: 83-117.
- Candlish, Stewart. "The Truth about F.H. Bradley". *Mind* 98 (1989): 331-348.
- David, Marian. *Correspondence and Disquotation*. New York: Oxford University Press,  
1994.
- Davidson, Donald. "A Coherence Theory of Truth and Knowledge", in *Truth and Interpretation, Perspectives on the Philosophy of Donald Davidson*, Ernest Lepore (ed.), Oxford: Basil Blackwell, 1984: 307-19.
- Davidson, Donald. "The Folly of Trying to Define Truth". *Journal of Philosophy* 93 (1996): 263-279; repr. in Blackburn and Simmons (eds.) 1999: 308-322.
- Devitt, Michael. *Realism and Truth*, 2<sup>nd</sup> edition. Oxford: Blackwell, 1991.
- Dewey, John. *Essays in Experimental Logic*. Chicago: University of Chicago Press,  
1916.
- Dewey, John. *Logic: The Theory of Inquiry*. New York: Holt, Rinehart and Winston,  
1938.
- Dodd, Julian. *An Identity Theory of Truth*. London: Macmillan, 2000.
- Dummett, Michael. *Truth and Other Enigmas*. Cambridge, Mass.: Harvard University Press, 1978.
- Etchemendy, Jon. "Tarski on Truth and Logical Consequence". *Journal of Symbolic Logic* 53 (1988): 51-79.
- Field, Hartry. "Tarski's Theory of Truth". *Journal of Philosophy* 69 (1972): 347-375.
- Field, Hartry. "Deflationist views of meaning and content". *Mind* 103 (1994): 249-84;  
repr. in Blackburn and Simmons (eds.) 1999: 351-391.

- Frege, Gottlob. "Der Gedanke: eine logische Untersuchung" (1918), translated as "The Thought: A Logical Inquiry" in *Mind* 65 (1956): 289-311; repr. in Blackburn and Simmons (eds.) 1999: 85-105.
- Grover, D., Camp, J., and Belnap, N. "A Prosentential Theory of Truth". *Philosophical Studies* 27 (1975): 73-125.
- Gupta, Anil. "Minimalism". In James E. Tomberlin (ed.), *Philosophical Perspectives* vii, *Language and Logic*. Atascadero: Ridgeview Press, 1993: 359-369.
- Gupta, Anil, and Belnap, Nuel. *The Revision Theory of Truth*. Cambridge, Mass.: MIT Press, 1993.
- Hill, Christopher. *Thought and World*. Cambridge: Cambridge University Press, 2002.
- Hornsby, Jennifer. "Truth: The Identity Theory", *Proceedings of the Aristotelian Society* 97 (1997): 1-24.
- Horwich, Paul. *Truth*. Oxford: Basil Blackwell, 1990. 2<sup>nd</sup> edition, Oxford: Oxford University Press, 1998.
- James, William. *Pragmatism*. New York: Longmans, Green and Co., 1907.
- James, William. *The Meaning of Truth*. New York: Longmans, Green and Co., 1909.
- Joachim, H. H. *The Nature of Truth*. Oxford: Oxford University Press, 1906
- Kirkham, Richard. *Theories of Truth*. Cambridge, Mass.: MIT Press, 1992.
- Kripke, Saul. "Outline of a Theory of Truth". *Journal of Philosophy* 72 (1975): repr. in Martin (ed.) 1984: 53-81.
- Martin, Robert L. *Recent Essays on Truth and the Liar Paradox*. Oxford: Oxford University Press, 1984.
- Moore, G. E. 1910-11 lectures published as *Some Main Problems of Philosophy*. New York: Collier Books, 1953.

- Neurath, Otto. "Protokollsätze". *Erkenntnis* 3 (1932-33). Tr. as "Protocol Sentences", in *Logical Positivism*, ed. by A. J. Ayer, Glencoe, Ill.: The Free Press, 1959:199-208.
- Peirce, Charles. "The Fixation of Belief". *Popular Science Monthly* 1877. Repr. in *Philosophical Writings of Peirce*, selected and edited with an introduction by Justus Buchler, New York: Dover Publications 1955: 5-22.
- Peirce, Charles. "How to Make Our Ideas Clear". *Popular Science Monthly* 1878. Repr. in *Philosophical Writings of Peirce*, selected and edited with an introduction by Justus Buchler, New York: Dover Publications 1955: 23-41.
- Plato. *Sophist*. Translated by F. M. Cornford. In *Plato: The Collected Dialogues*, edited by Edith Hamilton and Huntington Cairns. Princeton, N.J.: Princeton University Press, 1961.
- Priest, Graham. *In Contradiction: A Study of the Transconsistent*. The Hague: Nijhoff, 1987.
- Priest, Graham, Beall, J.C., and Armour-Garb, Brad. *The Law of Non-Contradiction: New Philosophical Essays*. Oxford: Oxford University Press, 2004.
- Putnam, Hilary. *Reason, Truth and History*. Cambridge: Cambridge University Press, 1981.
- Quine, W.V. *Philosophy of Logic*. Englewood Cliffs: Prentice Hall, 1970.
- Ramsey, Frank. "Facts and Propositions". *Proceedings of the Aristotelian Society* suppl.vol. 7 (1927): 153-71. Extracted in Blackburn and Simmons (eds.) 1999: 106-107; page references are to this extract.
- Rorty, Richard. *Consequences of Pragmatism*. Minneapolis: University of Minnesota Press, 1982.



- Russell, Bertrand. "On the Nature of Truth". *Proceedings of the Aristotelian Society* 7 (1906-7): 28-49.
- Russell, Bertrand. "Mathematical logic as based on the theory of types". *American Journal of Mathematics* 30 (1908): 222-62. Repr. in Jean van Heijenoort (ed.), *From Frege to Gödel*, Cambridge, Mass.: Harvard University Press, 1967: 150-182.
- Russell, Bertrand. *The Problems of Philosophy*. London: Home University Press 1912. repr. at Oxford: Oxford University Press 1959.
- Russell, Bertrand. *The Philosophy of Logical Atomism*, 1918 lectures. In *Logic and Knowledge*, edited by R. C. Marsh. London: George Allen and Unwin, 1956.
- Schlick, Moritz. "Ueber das Fundament der Erkenntnis". *Erkenntnis* 4 (1934). Tr. as "The Foundation of Knowledge", in *Logical Positivism*, ed. by A. J. Ayer, Glencoe, Ill.: The Free Press, 1959: 209-227.
- Simmons, Keith. *Universality and the Liar*. Cambridge: Cambridge University Press, 1993.
- Strawson, P. F. "Truth". *Analysis* 9 (1949): 83-97.
- Tarski, Alfred. "The Concept of Truth in Formalized Languages", first published in Polish 1930-1931. In *Logic, Semantics, Metamathematics* (2<sup>nd</sup> edition), trans. J. H. Woodger. Indianapolis: Hackett, 1983.
- Tarski, Alfred. "The Semantic Conception of Truth and the Foundations of Semantics", *Philosophical and Phenomenological Research* 4 (1944): 342-360. Repr. in Blackburn and Simmons (eds.) 1999: 115-143; page references are to this reprint.
- Wittgenstein, Ludwig. *Tractatus Logico-Philosophicus*. First published in German in 1921. English translation at London: Routledge & Kegan Paul, 1922.

Truth

Keith Simmons

Author: Keith Simmons

Word count: 5688 words.