# TRUTH

## VIRTUAL ISSUE NO. 1

*Featuring classic papers from the archive &
commentaries by contemporary philosophers*

# TRUTH

F.P. RAMSEY / PETER SULLIVAN
J.L. AUSTIN / CHARLES TRAVIS
P.F. STRAWSON / PAUL SNOWDON
MICHAEL DUMMETT / IAN RUMFITT
PETER GEACH / GRAHAM OPPY
JANE HEAL / JULIAN DODD
JENNIFER HORNSBY / GILA SHER

## GUEST EDITOR
*Guy Longworth, University of Warwick*

# CONTENTS

# INTRODUCTION

# INTRODUCTION

## GUY LONGWORTH

IN celebration of the 125th year of the Proceedings, we are proud to present the first *Virtual Issue of the Proceedings of the Aristotelian Society*. The *Virtual Issue* is based upon an Online Conference on the theme of Truth that took place 12th–18th April 2013. This weeklong event featured papers from our back catalogue, commentaries on these papers delivered by contemporary philosophers, and an online-based discussion forum that was open to all. The *Virtual Issue* comprises the classic papers and commentaries from the conference.

Questions about truth have figured centrally in Philosophy throughout its history. What is it for the things we say or believe to be true? Does truth depend on a relation between what we say or believe and the world? What are the natures of the things we say or believe, the bearers of truth? To what are the truth-bearers related when they are true: are they related to facts, ordinary objects, or something else? What is the required relation? We'll want an account of the nature of truth that addresses those questions also to fit with an account of truth's importance: why should it matter to us that what we say or believe is true rather than false? Our views about truth are liable to impact widely on our views about other things. Are moral claims or views apt to be true or false, or are they to be evaluated along different dimensions? Does truth figure in an account of the nature of belief or the nature of assertion? Is the acquisition of beliefs that are true amongst the fundamental aims of inquiry?

The papers selected for the Online Conference and *Virtual Issue* were chosen for the distinctive answers that they advance to some or all of these questions. In some cases, papers were chosen because they have had a decisive impact on later discussions. In some cases, papers were chosen because they present views and arguments that deserve more careful consideration than they have thus far received. In all cases, there is much to be gained from becoming acquainted, or reacquainted, with these important texts. The main aim of the commentaries is to stimulate discussion by highlighting major themes in its associated paper and pointing to ways in which those themes are of continuing importance to current debates. Some of them also point to specific challenges that

might be brought against claims or arguments in the associated paper and indicate connections with themes discussed in the other papers.

The remainder of this introduction to the conference theme presents a slightly more detailed overview of some of the central philosophical questions about truth that are discussed within the target papers and commentaries. It also provides some links to useful entries in the *Stanford Encyclopedia of Philosophy.*

In its most general form, the central philosophical question about truth can be stated simply. We say, and judge, that various things are true or false. For example, suppose that you think that snow is white. I might judge that what you thereby think is true. What would it take for my judgment to be correct, for what you think to be true? What is the most fundamental account of what it is for the sorts of things that can be true to be true?

More specifically, we can consider the following two questions:

1. What do we say to be true or false? Is it things people think? If it is, what are the things people think? Is it things people say? Is it sentences that people use to say things? Is it episodes in which people say things or think things?

2. What is required for the things we say to be true to be true? On one view, for something to be true it must correspond with the world, or with the way things are. This is the general form taken by correspondence theories of truth. If a theory of that general form is right, further questions arise. First, what must something correspond with if it is to be true? Must it correspond with the facts? For example, is the claim that snow is white true because it corresponds with the facts? Or is it true because it corresponds more specifically with the fact that snow is white? If it is the facts, what are they? And if it is the fact that snow is white, what is the nature of that fact? A second range of questions that arise concern the nature of correspondence. Does correspondence amount to some sort of similarity between the things that are true and the things that make them true—the truth-makers? Or is the relation more intimate than that? Are truths identical with facts? Almost all of the pieces discuss correspondence theories of truth. Hornsby's piece defends an identity theory of truth, on which truths are identified with facts.

Many philosophers have held that truth depends upon one or another form of correspondence between things that are true and other things. However, many other philosophers think that such a view is mistaken. The most radical amongst the latter group reject correspondence theories of truth because they hold that there is really nothing very much to be said about truth. Such philosophers endorse versions of deflationary theories of truth. More specifically, defenders of deflationary theories of truth focus on what many people take to be a platitude about truth, that claims of the following forms are bound to be correct:

> (S) The sentence 'S' is true if and only if S.

> For example, the sentence 'Snow is white' is true if and only if snow is white.

> (P) The claim, thought, assertion, or statement that P is true if and only if P.

> For example, the statement that snow is white is true if and only if snow is white.

Now many philosophers, including many defenders of correspondence theories of truth, agree that such claims are bound to be true. However, those that defend correspondence theories think that agreeing to that much is consistent with saying more about the nature of truth. By contrast, defenders of deflationary theories of truth hold (roughly) that there is no more to be said about truth than that claims of the form (S) or (P) are bound to be true. The nature of truth, insofar as it has a nature, is fully captured by its role in guaranteeing the truth of claims like (S) or (P). Ramsey's piece provided inspiration for deflationary theories of truth. Such theories are explicitly discussed in, or figure in the background of, all of the pieces.

One reason that addressing such questions about truth is important to us is that truth itself seems important to us. It seems important to us that our claims and beliefs are correct, and that seems to depend in turn on whether what we claim and believe is true. In that sense, we seem to value truth. That fact about our attitude towards truth raises further questions. First, if truth really is valuable, why is it valuable? What is it about truths, as opposed to falsehoods, that makes them distinctively valuable to us? One option for answering this question would be to deny that truth is distinctively valuable. Second, what does the claim that

truth is valuable amount to? Does it amount to the claim that we ought always to seek out truths and avoid falsehoods? If it does, could that demand on us be trumped by other demands? Should one never believe anything false even if believing it can help us to get things done? For example, should one have no general beliefs about the physical world rather than believing the false claims made in Newtonian Mechanics? And should one always seek out truths, even where those truths are useless or uninteresting? For example, should one aim to count one's books merely so that one can acquire a true belief concerning their number? Such questions about the value of truth play central roles in the pieces by Dummett, Geach, and Heal. They also figure in the background of all of the pieces, because any account of the nature of truth will need to connect with an account of truth's value.

<div align="center">*</div>

Here are some links to useful articles in the *Stanford Encyclopedia of Philosophy*:

Michael Glanzberg's general entry on truth:

http://plato.stanford.edu/entries/truth/

Daniel Stoljar and Nic Damnjanovic's entry on deflationary theories of truth (especially relevant to Ramsey, Austin, Strawson, Dummett, Geach, Heal, Hornsby):

http://plato.stanford.edu/entries/truth-deflationary/

Marian David's entry on correspondence theories of truth (especially relevant to Ramsey, Austin, Dummett, Geach, Hornsby):

http://plato.stanford.edu/entries/truth-correspondence/

Fraser McBride's entry on truth-makers (especially relevant to Ramsey, Strawson, Austin, Dummett, Geach, Hornsby):

http://plato.stanford.edu/entries/truthmakers/

Stewart Candlish and Nic Damnjanovic's entry on the identity theory of truth (especially relevant to Hornsby):

http://plato.stanford.edu/entries/truth-identity/

Guy Longworth's entry on J. L. Austin (especially relevant to Austin, Strawson):

http://plato.stanford.edu/entries/austin-jl/

Paul Snowdon's entry on P. F. Strawson (especially relevant to Austin, Strawson):

http://plato.stanford.edu/entries/strawson/

Alexander Miller's entry on realism (especially relevant to Dummett):

http://plato.stanford.edu/entries/realism/

F.P. RAMSEY & PETER SULLIVAN

# BIOGRAPHY

Frank Plumpton Ramsey (22 February 1903 – 19 January 1930) was a British mathematician who, in addition to mathematics, made significant contributions in philosophy and economics before his death at the age of 26. He was a close friend of Ludwig Wittgenstein, and was instrumental in translating Wittgenstein's *Tractatus Logico-Philosophicus* into English, and in persuading Wittgenstein to return to philosophy and Cambridge.

"Facts and Propositions" was originally published in the *Proceedings of the Aristotelian Society, Supplementary Volume VII* (1927).

Peter Sullivan's research interests are in the history of analytic philosophy and the philosophy of logic and mathematics. The principal focus of much of his research has been Wittgenstein's *Tractatus Logico-Philosophicus*; surrounding this work, he has published papers on Frege, Russell, and Ramsey.

From 2003-2006, Peter was involved in an AHRC-funded research project on the interpretation of the *Tractatus* which aimed to explore how Wittgenstein's critical conception of the task and method of philosophy emerges out of his detailed engagement with the logical theories of his predecessors in the analytic tradition. Since then Peter has been working particularly on Ramsey, both in his own right and as a resource for understanding Wittgenstein.

# Facts and Propositions

## F.P. RAMSEY

# FACTS AND PROPOSITIONS

## F.P RAMSEY

THE problem with which I propose to deal is the logical analysis of what may be called by any of the terms judgment, belief, or assertion. Suppose I am at this moment judging that Caesar was murdered; then it is natural to distinguish in this fact on the one side either my mind, or my present mental state, or words or images in my mind, which we will call the mental factor or factors, and on the other side either Caesar or Caesar's murder, or Caesar and murder, or the proposition Caesar was murdered, or the fact that Caesar was murdered, which we will call the objective factor or factors, and to suppose that the fact that I am judging that Caesar was murdered consists in the holding of some relation or relations between these mental and objective factors. The questions that arise are in regard to the nature of the two sets of factors and of the relations between them, the fundamental distinction between these elements being hardly open to question.

Let us begin with the objective factor or factors; the simplest view is that there is one such factor only, a proposition, which may be either true or false, truth and falsity being unanalysable attributes. This was at one time the view of Mr. Russell, and in his essay, "On the Nature of Truth and Falsehood,"[1] he explains the reasons which led him to abandon it. These were, in brief, the incredibility of the existence of such objects as "that Caesar died in his bed," which could be described as objective falsehoods, and the mysterious nature of the difference, on this theory, between truth and falsehood. He therefore concluded, in my opinion rightly, that a judgment had no single object, but was a multiple relation of the mind or mental factors to many objects, those, namely, which we should ordinarily call constituents of the proposition judged.

There is, however, an alternative way of holding that a judgment has a single object, which it would be well to consider before we pass on. In the above-mentioned essay Mr. Russell asserts that a perception, which unlike judgment he regards as infallible, has a single object, for instance, the complex object "knife-to-left-of-book." This complex object can, I think, be identified with what many people (and Mr. Russell now) would call the *fact* that the knife is to the left of the book; we could, for

---

[1] In *Philosophical Essays*, 1910.

instance, say that we perceived this fact. And just as, if we take any true proposition such as that Caesar did not die in his bed, we can form a corresponding phrase beginning with " the fact that " and talk about the fact that he did not die in his bed, so Mr. Russell supposed that to any true proposition there corresponded a complex object.

Mr. Russell, then, held that the object of a perception was a fact, but that in the case of judgment the possibility of error made such a view untenable, since the object of a judgment that Caesar died in his bed could not be the fact that he died in his bed, as there was no such fact. It is, however, evident that this difficulty about error could be removed by postulating for the case of judgment two different relations between the mental factors and the fact, one occurring in true judgments, the other in false. Thus, a judgment that Caesar was murdered and a judgment that Caesar was not murdered would have the same object, the fact that Caesar was murdered, but differ in respect of the relations between the mental factor and this object. Thus, in the *Analysis of Mind*[2], Mr. Russell speaks of beliefs as either pointing towards or pointing away from facts. It seems to me, however, that any such view either of judgment or of perception would be inadequate for a reason, which, if valid, is of great importance. Let us for simplicity take the case of perception, and assuming for the sake of argument that it is infallible, consider whether "he perceives that the knife is to the left of the book" can really assert a dual relation between a person and a fact. Suppose that I who make the assertion cannot myself see the knife and book, that the knife is really to the right of the book; but that through some mistake I suppose that it is on the left and that he perceives it to be on the left, so that I assert falsely "he perceives that the knife is to the left of the book." Then my statement, though false, is significant, and has the same meaning as it would have if it were true; this meaning cannot therefore be that there is a dual relation between the person and something (a fact) of which "that the knife is to the left of the book" is the name, because there is no such thing. The situation is the same as that with descriptions; "the King of France is wise" is not nonsense, and so "the King of France," as Mr. Russell has shown, is not a name but an incomplete symbol, and the same must be true of "the King of Italy." So also "that the knife is to the left of the book," whether it is true or false, cannot be the name of a fact.

---

[2] P. 272 – It should be observed that in the *Analysis of Mind*, a "belief " is what we call a mental factor, not the whole complex mental factors-relations-objective factors.

But, it will be asked, why should it not be a description of a fact? If I say, "he perceives that the knife is to the left of the book," I mean that he perceives a fact, which is not named but described as of a certain sort, and the difficulty will disappear when my assertion is analysed according to Mr. Russell's theory of descriptions. Similarly, it will be said, "the death of Caesar" is a description of an event, and "the fact that Caesar died" is only an alternative expression for "the death of Caesar."

Such an objection is plausible but not, in my opinion, valid. The truth is that a phrase like "the death of Caesar" can be used in two different ways; ordinarily, we use it as the description of an event, and we could say that "the death of Caesar" and "the murder of Caesar" were two different descriptions of the same event. But we can also use "the death of Caesar" in a context like "he was aware of the death of Caesar " meaning "he was aware that Caesar had died"; here (and this is the sort of case which occurs in the discussion of cognition) we cannot regard "the death of Caesar" as the description of an event; if it were, the whole proposition would be, "There is an event E of a certain sort, such that he is aware of E," and would be still true if we substituted another description of the same event, e.g., "the murder of Caesar." That is, if his awareness has for its object an event described by "the death of Caesar," then, if he is aware of the death of Caesar, he must also be aware of the murder of Caesar, for they are identical. But, in fact, he could quite well be aware that Caesar had died, without knowing that he had been murdered, so that his awareness must have for its object not merely an event but an event and a character also.

The connection between the event which was the death of Caesar and the fact that Caesar died is, in my opinion, this: "That Caesar died" is really an existential proposition, asserting the existence of an event of a certain sort, thus resembling "Italy has a King," which asserts the existence of a man of a certain sort. The event which is of that sort is called the death of Caesar and must no more be confused with the fact that Caesar died, than the King of Italy should be confused with the fact that Italy has a King.

We have seen, then, that a phrase beginning "the fact that" is not a name, and also not a description; it is, therefore, neither a name nor a description of any genuine constituent of a proposition, and so a proposition about "the fact that $aRb$" must be analysed into (1) the proposition $aRb$, (2) some further pro-position about $a$, $R$, $b$, and other things; and an analysis of cognition in terms of relations to facts cannot

be accepted as ultimate. We are driven, therefore, to Mr. Russell's conclusion that a judgment[3] has not one object but many, to which the mental factor is multiply related; but to leave it at that, as he did, cannot be regarded as satisfactory. There is no reason to suppose the multiple relation simple, it may, for instance, result from the combination of dual relations between parts of the mental factor and the separate objects, and it is desirable that we should try to find out more about it, and how it varies when the form of proposition believed is varied. Similarly, a theory of descriptions which contented itself with observing that "the King of France is wise" could be regarded as asserting a possibly complex multiple relation between kingship, France, and wisdom, would be miserably inferior to Mr. Russell's theory, which explains exactly what relation it is.

But before we proceed further with the analysis of judgment, it is necessary to say something about truth and falsehood, in order to show that there is really no separate problem of truth but merely a linguistic muddle. Truth and falsity are ascribed primarily to propositions. The proposition to which they are ascribed may be either explicitly given or described. Suppose first that it is explicitly given; then it is evident that "it is true that Caesar was murdered" means no more than that Caesar was murdered, and "it is false that Caesar was murdered" means that Caesar was not murdered. They are phrases which we sometimes use for emphasis or for stylistic reasons, or to indicate the position occupied by the statement in our argument. So also we can say "it is a fact that he was murdered" or "that he was murdered is contrary to fact."

In the second case in which the proposition is described and not given explicitly, we have perhaps more of a problem, for we get statements from which we cannot in ordinary language eliminate the words "true" and "false." Thus if I say "he is always right" I mean that the propositions he asserts are always true, and there does not seem to be any way of expressing this without using the word "true." But suppose we put it thus "For all $p$, if he asserts $p$, $p$ is true," then we see that the propositional function $p$ is true is simply the same as $p$, as *e.g.* its value "Caesar was murdered is true," is the same as "Caesar was murdered." We have in English to add "is true" to give the sentence a verb, forgetting that "$p$" already contains a (variable) verb. This may perhaps be made clearer by supposing, for a moment, that only one form of proposition is in question, say the relational form $aRb$; then "he is always right" could be expressed by "For all $a$, $R$, $b$, if he asserts $aRb$,

---

[3] And, in our view, any other form of knowledge or opinion that something is the case.

then $aRb$" to which "is true" would be an obviously superfluous addition. When all forms of proposition are included the analysis is more complicated but not essentially different, and it is clear that the problem is not as to the nature of truth and falsehood, but as to the nature of judgment or assertion, for what is difficult to analyse in the above formulation is "he asserts $aRb$."

It is, perhaps, also immediately obvious that if we have analysed judgment we have solved the problem of truth; for taking the mental factor in a judgment (which is often itself called a judgment), the truth or falsity of this depends only on what proposition it is that is judged, and what we have to explain is the meaning of saying that the judgment is a judgment that $a$ has $R$ to $b$, *i.e.* is true if $aRb$, false if not. We can, if we like, say that it is true if there exists a corresponding fact that $a$ has $R$ to $b$, but this is essentially not an analysis but a periphrasis for "the fact that $a$ has $R$ to $b$ exists" is no different from "$a$ has $R$ to $b$."

In order to proceed further, we must now consider the mental factors in a belief. Their nature will depend on the sense in which we are using the ambiguous term belief: it is, for instance, possible to say that a chicken believes a certain sort of caterpillar to be poisonous, and mean by that merely that it abstains from eating such caterpillars on account of unpleasant experiences connected with them. The mental factors in such a belief would be parts of the chicken's behaviour, which are somehow related to the objective factors, viz., the kind of caterpillars and poisonousness. An exact analysis of this relation would be very difficult, but it might well be held that in regard to this kind of belief the pragmatist view was correct, *i.e.* that the relation between the chicken's behaviour and the objective factors was that the actions were such as to be useful if, and only if, the caterpillars were actually poisonous. Thus any actions for whose utility $p$ is a necessary and sufficient condition might be called a belief that $p$, and so would be true if $p$, i.e. if they are useful.[4]

But without wishing to depreciate the importance of this kind of belief, it is not what I wish to discuss here. I prefer to deal with those beliefs which are expressed in words, or possibly images or other symbols, consciously asserted or denied; for these, in my view, are the most proper subject for logical criticism.

---

[4] It is useful to believe $aRb$ would mean It is useful to do things which are useful if, and only if, $aRb$; which is evidently equivalent to $aRb$.

The mental factors of such a belief I take to be words, spoken aloud or to one's self or merely imagined, connected together and accompanied by a feeling or feelings of belief or disbelief, related to them in a way I do not propose to discuss.[5] I shall suppose for simplicity that the thinker with whom we are concerned uses a systematic language without irregularities and with an exact logical notation like that of *Principia Mathematica*. The primitive signs in such a language can be divided into names, logical constants, and variables. Let us begin with names; each name means an object, meaning being a dual relation between them. Evidently name, meaning, relation, and object may be really all complex, so that the fact that the name means the object is not ultimately of the dual relational form but far more complicated.[6] Nevertheless, just as in the study of chess, nothing is gained by discussing the atoms of which the chessmen are composed, so in the study of logic nothing is gained by entering into the ultimate analysis of names and the objects they signify. These form the elements of the thinker's beliefs, in terms of which the various logical relations of one belief to another can all be stated, and their internal constitution is immaterial.

By means of names alone the thinker can form what we may call atomic sentences, which from our formal standpoint offer no very serious problem. If $a$, $R$, and $b$ are things which are simple in relation to his language, *i.e.* of the types for instances of which he has names, he will believe that $aRb$ by having names for $a$, $R$, and $b$ connected in his mind and accompanied by a feeling of belief. This statement is, however, too simple since the names must be united in a way appropriate to $aRb$ rather than to $bRa$; this can be explained by saying that the name of R is not the word "R," but the relation we make between "$a$" and "$b$" by writing "$aRb$." The sense in which this relation unites "$a$" and "$b$," then determines whether it is a belief that $aRb$ or that $bRa$. There are various other difficulties of the same sort, but I propose to pass on to the more interesting problems which arise when we consider more complicated beliefs, which require for their expression not only names but logical constants as well, so that we have to explain the mode of significance of such words as "not " and "or."

---

[5] I speak throughout as if the differences between belief, disbelief, and mere consideration lay in the presence or absence of "feelings "; but any other word may be substituted for "feeling " which the reader prefers, *e.g.* "specific quality" or "act of assertion" and "act of denial."

[6] This is most obvious in the case of names, which generally consist of letters, so that their complexity is evident.

One possible explanation[7] is that they, or some of them, *e.g.* "not" and "and" in terms of which the others can be defined, are the names of relations, so that the sentences in which they occur are similar to atomic ones except that the relations they assert are logical instead of material. On this view every proposition is ultimately affirmative, asserting a simple relation between simple terms, or a simple quality of a simple term. Thus, "this is not-red" asserts a relation of negation between this and redness, and "this is not not-red" another relation of negation between this, redness and the first relation of negation.

This view requires such a different attitude to logic from mine that it is difficult for me to find a common basis from which to discuss it. There are, however, one or two things I should like to say in criticism – first, that I find it very unsatisfactory to be left with no explanation of formal logic: except that it is a collection of "necessary facts." The conclusion of a formal inference must, I feel, be in some sense contained in the premisses and not something new; I cannot believe that from one fact, *e.g.* that a thing is red, it should be possible to infer an infinite number of different facts, such as that it is not not-red, and that it is both red and not not-red. These, I should say, are simply the same fact expressed by other words; nor is it inevitable that there should be all these different ways of saying the same thing. We might, for instance, express negation not by inserting a word "not," but by writing what we negate upside down. Such a symbolism is only inconvenient because we are not trained to perceive complicated symmetry about a horizontal axis, and if we adopted it we should be rid of the redundant "not-not," for the result of negating the sentence "$p$" twice would be simply the sentence "$p$" itself.

It seems to me, therefore, that "not" cannot be a name (for if it were, "not-not-$p$" would have to be about the object not and so different in meaning from "$p$"), but must function in a radically different fashion. It follows that we must allow negations and disjunctions to be ultimately different from positive assertions and not merely the assertions of different but equally positive relationships. We must, therefore, abandon the idea that every proposition asserts a relation between terms, an idea that seems as difficult to discard as the older one that a proposition always asserted a predicate of a subject.

Suppose our thinker is considering a single atomic sentence, and that the progress of his meditation leads either to his believing it or his disbelieving it. These may be supposed to consist originally in two

---

[7] See, especially, J. A. Chadwick, "Logical Constants," *Mind*, Jan., 1927.

different feelings related to the atomic sentence, and in such a relation mutually exclusive; the difference between assertion and denial thus consisting in a difference of feeling and not in the absence or presence of a word like "not." Such a word will, however, be almost indispensable for purposes of communication, belief in the atomic sentence being communicated by uttering it aloud, disbelief by uttering it together with the word "not." By a sort of association this word will become part of the internal language of our thinker, and instead of feeling disbelief towards "$p$" he will sometimes feel belief towards "not-$p$."

If this happens we can say that disbelieving "$p$" and believing "not-$p$" are equivalent occurrences, but to determine what we mean by this "equivalent" is, to my mind, the central difficulty of the subject. The difficulty exists on any theory, but is particularly important on mine, which holds that the significance of "not" consists not in a meaning relation to an object, but in this equivalence between disbelieving "$p$" and believing "not-$p$."

It seems to me that the equivalence between believing "not-$p$" and disbelieving "$p$" is to be defined in terms of causation, the two occurrences having in common many of their causes and many of their effects. There would be many occasions on which we should expect one or other to occur, but not know which, and whichever occurred we should expect the same kind of behaviour in consequence. To be equivalent, we may say, is to have in common certain causal properties, which I wish I could define more precisely. Clearly they are not at all simple; there is no uniform action which believing "$p$" will always produce. It may lead to no action at all, except in particular circumstances, so that its causal properties will only express what effects result from it when certain other conditions are fulfilled. And, again, only certain sorts of causes and effects must be admitted; for instance, we are not concerned with the factors determining, and the results determined by, the rhythm of the words.

Feeling belief towards the words "not-$p$" and feeling disbelief towards the words "$p$" have then in common certain causal properties. I propose to express this fact by saying that the two occurrences express the same attitude, the attitude of disbelieving $p$ or believing not-$p$. On the other hand, feeling belief towards "$p$" has different causal properties and so expresses a different attitude, the attitude of believing $p$. It is evident that the importance of beliefs and disbeliefs lies not in their intrinsic nature but in their causal properties, *i.e.* their causes and more especially their effects. For why should I want to have a feeling of belief towards names "$a$," "$R$," and "$b$" when $aRb$, and of disbelief when not-

*aRb*, except because the effects of these feelings are more often satisfactory than those of the alternative ones.

If then I say about someone whose language I do not know "he is believing that not-*aRb*," I mean that there is occurring in his mind such a combination of a feeling and words as expresses the attitude of believing not-*aRb*, *i.e.*, has certain causal properties, which can *in this simple case*[8] be specified as those belonging to the combination of a feeling of disbelief and names for *a*, *R*, and *b*, or, in the case of one who uses the English language, to the combination of a feeling of belief, names for *a*, *R*, and *b*, and an odd number of "not"'s. Besides this, we can say that the causal properties are connected with *a*, *R*, and *b* in such a way that the only things which can have them must be composed of names for *a*, *R*, and *b*. (This is the doctrine that the meaning of a sentence must result from the meaning of the words in it.)

When we are dealing with one atomic proposition only, we are accustomed to leave to the theory of probability the intermediate attitudes of partial belief, and consider only the extremes of full belief and full disbelief. But when our thinker is concerned with several atomic propositions at once, the matter is more complicated, for we have to deal not only with completely definite attitudes, such as believing *p* and disbelieving *q*, but also with relatively indefinite attitudes, such as believing that either *p* or *q* is true, but not knowing which. Any such attitude can, however, be defined in terms of the truth-possibilities of atomic propositions with which it agrees and disagrees. Thus, if we have *n* atomic propositions, with regard to their truth and falsity there are $2^n$ mutually exclusive possibilities, and a possible attitude is given by taking any set of these and saying that it is one of this set which is in fact realised, not one of the remainder. Thus, to believe *p* or *q* is to express agreement with the possibilities *p* true and *q* true, *p* false and *q* true, *p* true and *q* false, and disagreement with the remaining possibility *p* false and *q* false. To say that feeling belief towards a sentence expresses such an attitude, is to say that it has certain causal properties which vary with the attitude, *i.e.* with which possibilities are knocked out and which, so to speak, are still left in. Very roughly the thinker will act in disregard of the possibilities rejected, but how to explain this accurately I do not know.

---

[8] In the more complicated cases treated below a similar specification seems to me impossible, except by reference to a particular language. There are ways in which it can apparently be done, but, I think, they are illusory.

In any ordinary language such an attitude can be expressed by a feeling of belief towards a complicated sentence formed out of the atomic sentences by logical conjunctions; which attitude it is, depending not on the feeling but on the form of the sentence. We can therefore say elliptically that the sentence expresses the attitude, and that the meaning of a sentence is agreement and disagreement with such and such truth-possibilities, meaning by that that one who asserts or believes the sentence so agrees and disagrees.

In most logical notations the meaning of the sentence is determined by logical operation signs that occur in it, such as "not" and "and." These mean in the following way : "not-$P$," whether "$P$" be atomic or not, expresses agreement with the possibilities with which "$P$" expresses disagreement and vice versa. "$P$ and $Q$" expresses agreement with such possibilities, as both "$P$" and "$Q$" express agreement with, and disagreement with all others. By these rules the meaning of any sentence constructed from atomic sentences by means of "not" and "and" is completely determined; the meaning of "not" being thus a law determining the attitude expressed by "not-$P$" in terms of that expressed by "$P$."

This could, of course, only be used as a *definition* of "not" in a symbolism based directly on the truth-possibilities. Thus in the notation explained on page 95 of Mr. Wittgenstein's *Tractatus Logico-Philosophicus*, we could define "not-$P$" as the symbol obtained by interchanging the T's and blanks in the last column of "$P$." Ordinarily, however, we always use a different sort of symbolism in which "not" is a primitive sign which cannot be defined without circularity ; but even in this symbolism we can ask how '"nicht" means not' is to be analysed, and it is this question which the above remarks are intended to answer. In our ordinary symbolism the truth-possibilities are most conveniently expressed as conjunctions of atomic propositions and their negatives, and any proposition will be expressible as a disjunction of the truth-possibilities with which it agrees.

If we apply the logical operations to atomic sentences in an indiscriminate manner, we shall sometimes obtain composite sentences which express no attitude of belief. Thus "$p$ or not-$p$" excludes no possibility and so expresses no attitude of belief at all. It should be regarded not as a significant sentence but a sort of degenerate case,[9] and is called by Mr. Wittgenstein a tautology. It can be added to any other

---

[9] In the mathematical sense in which two lines or two points form a degenerate conic.

sentence without altering its meaning, for "$q$: $p$ or not-$p$ " agrees with just the same possibilities as "$q$" The propositions of formal logic and pure mathematics are in this sense tautologies, and that is what is meant by calling them "necessary truths."

Similarly, "$p$ and not-$p$" excludes every possibility and expresses no possible attitude: it is called a *contradiction*.

In terms of these ideas we can explain what is meant by logical, mathematical, or formal inference or implication. The inference from "$p$" to "$q$" is formally guaranteed when "if $p$, then $q$" is a tautology, or when the truth-possibilities with which "$p$" agrees are contained among those with which "$q$" agrees. When this happens, it is always possible to express "$p$" in the form "$q$ and $r$," so that the conclusion "$q$" can be said to be already contained in the premiss.

Before passing on to the question of general propositions I must say something about an obvious difficulty. We supposed above that the meanings of the names in our thinker's language might be really complex, so that what was to him an atomic sentence might after translation into a more refined language appear as nothing of the sort. If this were so it might happen that some of the combinations of truth and falsity of his atomic propositions were really self-contradictory. This has actually been supposed to be the case with "blue" and "red," and Leibniz and Wittgenstein have regarded "this is both blue and red" as being really self-contradictory, the contradiction being concealed by defective analysis. Whatever may be thought of this hypothesis, it seems to me that formal logic is not concerned with it, but presupposes that all the truth-possibilities of atomic sentences are really possible, or at least treats them as being so. No one could say that the inference from "this is red" to "this is not blue" was formally guaranteed like the syllogism. If I may revert to the analogy of chess this assumption might perhaps be compared to the assumption that the chessmen are not so strongly magnetised as to render some positions on the board mechanically impossible, so that we need only consider the restrictions imposed by the rules of the game, and can disregard any others which might conceivably arise from the physical constitution of the men.

We have so far confined ourselves to atomic propositions and those derived from them by any finite number of truth-operations, and unless our account is to be hopelessly incomplete we must now say something about general propositions such as are expressed in English by means of the words "all" and "some," or in the notation of *Principia Mathematica* by apparent variables. About these I adopt the view of Mr.

Wittgenstein[10] that "for all $x$, $fx$" is to be regarded as equivalent to the logical product of all the values of "$fx$" *i.e.* to the combination $fx_1$ and $fx_2$ and $fx_3$ and . . ., and that "there is an $x$ such that $fx$" is similarly their logical sum. In connection with such symbols we can distinguish first the element of generality, which comes in in specifying the truth-arguments, which are not, as before, enumerated, but determined as all values of a certain propositional function; and, secondly, the truth-function element which is the logical product in the first case and the logical sum in the second.

What is novel about general propositions is simply the specification of the truth-arguments by a propositional function instead of by enumeration. Thus general propositions, just like molecular ones, express agreement and disagreement with the truth-possibilities of atomic propositions, but they do this in a different and more complicated way. Feeling belief towards "for all $x$, $fx$" has certain causal properties which we call its expressing agreement only with the possibility that all the values of $fx$ are true. For a symbol to have these causal properties it is not necessary, as it was before, for it to contain names for all the objects involved combined into the appropriate atomic sentences, but by a peculiar law of psychology it is sufficient for it to be constructed in the above way by means of a propositional function.

As before, this must not be regarded as an attempt to define "all" and "some," but only as a contribution to the analysis of " *I believe that all (or some).* "

This view of general propositions has the great advantage that it enables us to extend to them Mr. Wittgenstein's account of logical inference, and his view that formal logic consists of tautologies. It is also the only view which explains how "$fa$" can be inferred from "for all $x$, $fx$," and "there is an $x$ such that $fx$ " from $fa$. The alternative theory that "there is an $x$ such that $fx$" should be regarded as an atomic proposition of the form "$F(f)$" ($f$ has application) leaves this entirely obscure; it gives no intelligible connection between $a$ being red and red having application, but abandoning any hope of explaining this relation is content merely to label it "necessary."

Nevertheless, I anticipate that objection will be made on the following lines: firstly, it will be said that $a$ cannot enter into the meaning of "for all $x$, $fx$," because I can assert this without ever having heard of $a$. To this I answer that this is an essential part of the utility of

---

[10] And also, apparently, of Mr. Johnson. See his *Logic*, Part II, p. 59.

the symbolism of generality, that it enables us to make assertions about things we have never heard of and so have no names for. Besides, that a is involved in the meaning of "for all x, fx" can be seen from the fact that if I say "for all x, fx," and someone replies "not-fa," then even though I had not before heard of a, he would undoubtedly be contradicting me.

The second objection that will be made is more serious; it will be said that this view of general propositions makes what things there are in the world not, as it really is, a contingent fact, but something presupposed by logic or at best a proposition of logic. Thus it will be urged that even if I could have a list of everything in the world "a," "b," ... ","  "for all x, fx" would still not be equivalent to "fa, fb... fz," but rather to "fa, fb... fz and a, b... z are everything." To this Mr. Wittgenstein would reply that "a, b... z are everything" is nonsense, and could not be written at all in his improved symbolism for identity. A proper discussion of this answer would involve the whole of his philosophy, and is, therefore, out of the question here; all that I propose to do is to retort with a *tu quoque!* The objection would evidently have no force if "a, b... z are everything" were, as with suitable definitions I think it can be made to be, a tautology; for then it could be left out without altering the meaning. The objectors will therefore claim that it is not a tautology, or in their terminology not a necessary proposition; and this they will presumably hold with regard to any proposition of the sort, *i.e.* they will say that to assert of a set of things that they are or are not everything cannot be either necessarily true or necessarily false. But they will, I conceive, admit that numerical identity and difference are necessary relations, that "there is an x such that fx" necessarily follows from "fa," and that whatever follows necessarily from a necessary truth is itself necessary. If so, their position cannot be maintained; for suppose a, b, c are, in fact, not everything, but that there is another thing d. Then that d is not identical with a, b, or c is a necessary fact; therefore it is necessary that there is an x, such that x is not identical with a, b, or c, or that a, b, c are not the only things in the world. This is, therefore, even on the objector's view, a necessary and not a contingent truth.

In conclusion, I must emphasise my indebtedness to Mr. Wittgenstein, from whom my view of logic is derived. Everything that I have said is due to him, except the parts which have a pragmatist tendency,[11] which seem to me to be needed in order to fill up a gap in his

---

[11]  And the suggestion that the notion of an atomic proposition may be relative to a language.

system. But whatever may be thought of these additions of mine, and however this gap should be filled in, his conception of formal logic seems to me indubitably an enormous advance on that of any previous thinker.

My pragmatism is derived from Mr. Russell; and is, of course, very vague and undeveloped. The essence of pragmatism I take to be this, that the meaning of a sentence is to be defined by reference to the actions to which asserting it would lead, or, more vaguely still, by its possible causes and effects. Of this I feel certain, but of nothing more definite.

*An Introduction to*
*'Facts and Propositions'*

PETER SULLIVAN

# AN INTRODUCTION TO
## 'FACTS AND PROPOSITIONS'

### PETER SULLIVAN


## I

FRANK Ramsey's 'Facts and Propositions', published in the Society's *Supplementary Volume* for 1927, is renowned as the inspiration for two highly influential and still actively debated views, one in the theory of truth, and one in the theory of content. In the theory of truth, this paper directly inspired the 'redundancy theory', which holds – in a sense that can be made more precise in various ways – that predication of truth makes no substantial addition to a language or system of thought from which it is absent; it thereby indirectly inspired various forms of 'disquotationalism' or 'minimalism' about truth, positions which aim to preserve much of the spirit of the redundancy theory while departing in various ways from its letter. In the theory of content the paper inspired a 'pragmatist' approach which seeks to explain what it is for a belief or other attitude, or an utterance expressing the attitude, to have a certain content by reference to its causal role. The most developed and explicitly Ramseyan form of this approach (Mellor 2012) is now known as 'success semantics': this theory identifies the content of a belief with the circumstance in which actions that would be caused by the belief, in conjunction various desires, would be successful in achieving the aims of those desires. As Blackburn nicely summarizes the approach, 'the idea is that we get our way, or flourish, … because we get things right about the world. The contents of our sentences [or beliefs] are then whatever it is that we get right' (2005, p. 22).

While it is completely clear that Ramsey inspired both of these theories, it is very much less clear whether he actually held either of them. The issues that would have to be resolved to settle either of these exegetical questions are complex, much more complex than can sensibly be addressed in a short introduction. What I attempt here is only to lay out two broad options for understanding the essay, and to mention some of the problems they face.

## II

A first approach to these exegetical questions emphasizes the issue of how Ramsey conceived his proposals in 'Facts and Propositions' as relating to the work of his predecessors. The essay opens with a question about the 'logical analysis' of belief, posed very much in the manner of Russell, who had spoken of 'the problem of the logical form of belief, i.e. what is the schema representing what occurs when a man believes' (1922, p. 19); its first section then endorses and extends Russell's argument for an answer to this question in broad accord with Russell's 'multiple-relation theory' of judgement. The essay ends with a very generous acknowledgement to Wittgenstein: 'Everything that I have said is due to him, except the parts that have a pragmatist tendency, which seem to me to be needed to fill up a gap in his system' (p. 170). Taking these facts as a guide to Ramsey's intentions, we should expect to find in the essay a largely Wittgensteinian answer to a Russellian problem. Speaking very broadly, it will straightaway seem hard to square this expectation with attribution to Ramsey of the theories he inspired. It is hard, that is, to see how a redundancy theory of truth might represent a view 'due to' Wittgenstein, or how success semantics might be designed to 'fill up a gap' in Wittgenstein's theory of content, the picture theory of the *Tractatus*.

A second approach to these questions emphasizes instead the need to account for how Ramsey conceived the relations amongst the various proposals advanced in 'Facts and Propositions'. Under the guiding expectation sketched in the previous paragraph, this would be a matter of understanding, first, how his brief remarks on the problem of truth are to serve as a corrective to Russell's theories, and secondly how the pragmatist elements Ramsey introduces can cohere with and supplement a basically Tractarian approach to the theory of content. But we should for the moment suspend that expectation, to bring into view the more general issue of what relation Ramsey takes there to be between the notions of truth, the topic of the first theory he inspired, and content, the topic of the second. Wittgenstein had held that to grasp the content of a sentence is 'to know what is the case if it is true' (1913, p. 104; cf. 1922, 4.063). Michael Dummett famously argued (1959) that this truth-conditional theory of content cannot be combined with a redundancy theory of truth: according to a redundancy theory, the content of a statement, 'It is true that $p$', is explained as being the same as that of the statement '$p$'; so the content of '$p$' cannot without circularity be

explained by that of 'It is true that *p*'.  If we suppose that Ramsey anticipated Dummett's argument, then attributing to him both of the theories he inspired would seem to yield a coherent schematic view of his aims in 'Facts and Propositions'.  According to this view Ramsey will be understood as endorsing the redundancy theory, and as recognizing that this endorsement requires a new, non-Wittgensteinian theory of content, one that explains content without appeal to the notion of truth; success semantics will then be understood as Ramsey's proposal to meet this need.  At the schematic level, then, this view offers us an intelligible account of what Ramsey was up to in the essay.  The difficulty this time is that it is hard – *very* hard – to see how this schema might connect with the detail of what Ramsey actually says.


III


Constraints of space mean that I can only illustrate some of these difficulties; and, given the focus of this on-line conference, the most relevant are those facing the suggestion that Ramsey advanced a redundancy theory of truth.

The first claim of a redundancy theory is that explicit predication of truth is redundant – that nothing can be said with it that is not at least as well said without it.  This much Ramsey clearly holds.  The sentence, 'It is true that Caesar was murdered', he says, 'means no more than that Caesar was murdered' (p. 157).  Some disquotationalists have held that a truth-predicate is eliminable in this kind of way only if the proposition to which it is ascribed is spelled out, and that the usefulness of the predicate shows itself when it is ascribed to propositions that are merely described or spoken of generally, as in 'Everything he says is true'. Ramsey goes further, holding that 'true' is eliminable even from these contexts.  In 'For all *p*, if he asserts *p*, *p* is true', he says, the need for the verb-phrase 'is true' is only an imposition of English grammar, which does not cope readily with generalization into sentence positions (p. 158).

But as Wittgenstein remarked, to do away with the words 'true' and 'false' is not to 'do away with the puzzles connected with truth and falsity' (1979a, p. 106).  The redundancy theorist is distinguished, not merely by advancing the eliminability claim, but by the explanation he offers for it:  he holds that the word 'true' is eliminable without loss because there is no substantial notion for it to express.  An alternative suggestion, due to Frege (1984, p. 354), is that explicit predication of

'true' *adds* nothing because the notion it expresses is *already present* in that to which it is ascribed.

Before we count Ramsey a redundancy theorist, then, we need to attend to more than the two paragraphs at pp. 157-8 in which he presents the eliminability claim: we need to attend at least to the engagement with Russell that leads up to this claim, and to the consequences Ramsey draws from it.

As to the latter, perhaps the most immediately telling fact is that, whereas a redundancy theorist might claim that there is really no problem about truth, Ramsey instead holds that 'there is really no *separate* problem of truth' (p. 157, emphasis added). That is, there is no problem of truth that can be separated from the question of what it is for a judgement to have a certain content. In Ramsey's argument the immediate role of the eliminability claim is to highlight this. It serves, for instance, to convert 'His judgement is true' into 'If he asserts $p$, then $p$', and plainly 'what is difficult to analyse in [this] formulation is "he asserts $p$"' (p. 158, variable altered).

A second telling fact is that in reformulating this point Ramsey shows no sign of shying away from, but instead seems clearly to endorse, the Tractarian view that for a judgement to have a certain content is for it to have a certain truth condition:

> It is, perhaps, also immediately obvious that if we have analysed judgement we have solved the problem of truth; for … the truth or falsity of [a judgement] depends only on what proposition it is that is judged, and what we have to explain is the meaning of saying that the judgement is a judgement that $a$ has $R$ to $b$, *i.e.* is true if $aRb$, false if not. (p. 158)

This should not be surprising, since in this paragraph Ramsey is simply repeating what he had said four years earlier, in exposition of Wittgenstein, in his Critical Notice of the *Tractatus* (1923, p. 275); in his lectures, too, Ramsey repeatedly presented instances of the eliminability claim as illustrations of Wittgenstein's analyses.

A third point worth noting is that Ramsey's eliminative paraphrases make use of a variable '$p$' in sentence position (one 'ranging over propositions'). Now Ramsey would be the last person simply to help himself to a variable without considering the question, what determines its range, or what notions are involved in grasping the generality it

expresses. The whole subsequent argument of 'Facts and Propositions' is structured by Ramsey's Tractarian answer to this question: the variable ranges over all truth-functions of elementary propositions.

If we turn now to the engagement with Russell that leads up to these two paragraphs, we find that Ramsey's central aim is to deny any explanatory role, in the analysis of cognition and hence in the analysis of truth, to the notion of a 'fact'. In adopting his multiple-relation theory of judgement Russell (1910) had come to hold that a judgement's having a certain content is not to be explained by its being coordinated with any unitary, complex entity – the kind of entity he had earlier called a 'proposition'. Russell however continued to hold that a judgement's being true *is* to be explained in this fashion: the truth of a judgement, he maintains, consists in the existence of a complex entity, a fact, corresponding to it. Ramsey's argument uses Russellian tools to unpick this Russellian position, and leads to the conclusion (p. 156) that a phrase of the form 'the fact that so-and-so' is no more a name or a description of a kind of complex entity than is a phrase of the form 'the proposition that so-and-so': *both* of these are – as Russell had said about the second but not about the first – 'incomplete symbols'. Again, in advancing this view (though not in every detail of his argument) Ramsey is following Wittgenstein, who in the *Notes on Logic* had declared, 'Neither the sense [truth-condition] nor the meaning of a proposition [the fact corresponding to it] is a thing. These words are incomplete symbols' (1913, p. 102). And again, the view is one that Ramsey first advanced in exposition of Wittgenstein (1923, p. 273).

Of course, these brief observations fall a long way short of a proof. But so far as they concern the first of the theories Ramsey inspired, the redundancy theory of truth, they do rather strongly suggest the following conclusions:

> (i)　　Ramsey's remarks on truth are presented in criticism of a correspondence theory of truth, as advanced by Russell from 1910;

> (ii)　　in presenting this criticism, Ramsey took himself to be developing views already advanced by Wittgenstein;

> (iii)　　in taking himself to agree with Wittgenstein on these matters, Ramsey was entirely right;

(iv)     Ramsey was therefore no more a redundancy theorist than was Wittgenstein.

Further, in suggesting these conclusions, the observations tend to favour the first of the two exegetical approaches distinguished in §II.
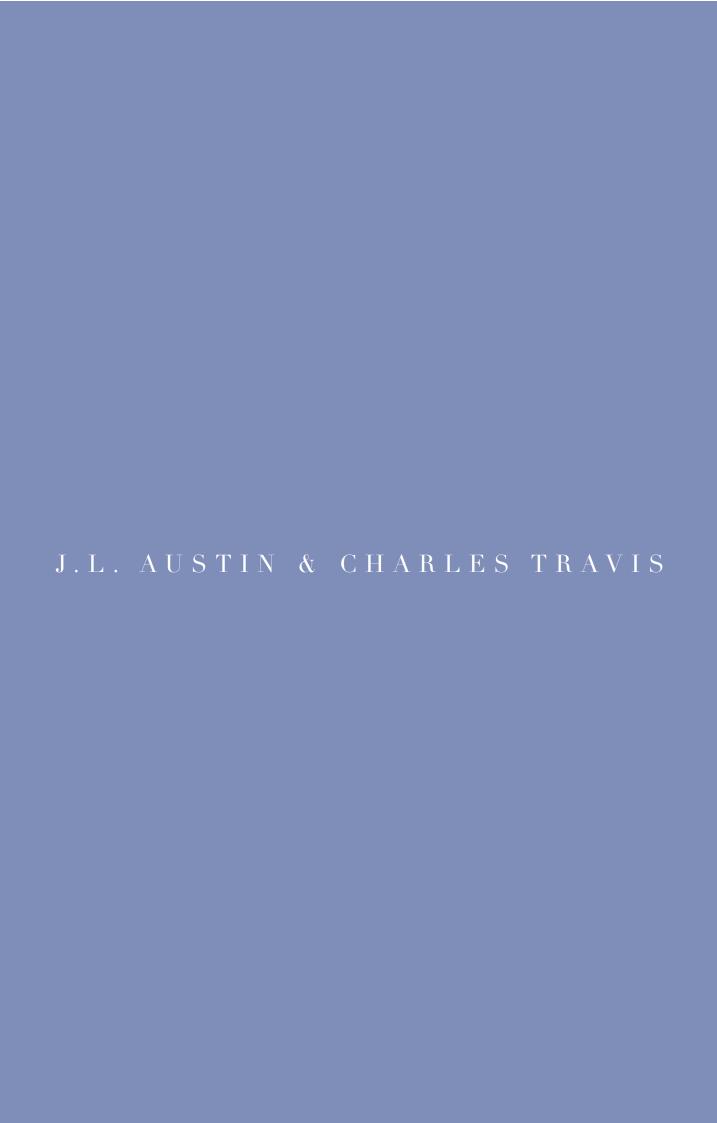

IV

Similar problems confront the attribution to Ramsey of the second theory his essay inspired, the theory of success semantics. The most obvious, of course, is that the example this theory develops – the example of a chicken believing that such-and-such caterpillars are poisonous – is offered by Ramsey as illustrating a 'kind of belief [that] is *not* what I want to discuss here' (p. 159, emphasis added). But much more important than this, to my mind, is that a reading in accordance with success semantics simply obliterates the compositional structure in beliefs by which Ramsey organizes his argument from p. 160 to the end of the essay. The 'gap' in Wittgenstein's theory, which Ramsey had already identified in his Critical Notice (1923, pp. 275-7), has to do with how the picture theory is to be extended from elementary to logically complex propositions. Ramsey's pragmatist proposals – centrally, his proposal that the meaning of 'not' is to be explained by the causal equivalence between the attitudes of believing 'not-*p*' and disbelieving '*p*' – are designed to fill precisely this gap. Success semantics, by contrast, specifies a condition for something to have a certain content that is independent of attributing any compositional structure to it. The two approaches could hardly be more different.

We should not, of course, confuse substantive and exegetical questions. The fact that Ramsey held a certain theory may be for some – it is, for instance, for me – a reason to take that theory very seriously. But for no one can the fact that Ramsey did not hold a certain theory be a reason to discount it. Success semantics, and the redundancy theory of truth, will stand or fall on their merits.

Returning, though, to exegetical questions, I hope these few introductory remarks will encourage readers towards the first broad approach distinguished in §II, and to try to understand 'Facts and Propositions' as offering a largely Wittgensteinian answer to a Russellian question. This approach faces serious problems of its own. But I do think they are the right problems.

REFERENCES

Blackburn, S. (2005), 'Success Semantics', in H. Lillehammer and D. H. Mellor, eds., *Ramsey's Legacy*.  Oxford: Oxford University Press.

Dummett, M. (1959), 'Truth', reprinted in his *Truth and Other Enigmas*. London: Duckworth, 1978.

Frege, G. (1984), *Collected Papers on Mathematics, Logic, and Philosophy*, ed. B. F. McGuiness.  Oxford: Blackwell.

Mellor, D. H. (2012), 'Successful Semantics', in his *Mind, Meaning, and Reality*. Oxford: Oxford University Press.

Ramsey, F. P. (1923), 'Critical Notice' of Wittgenstein (1922), reprinted in his *The Foundations of Mathematics and Other Logical Essays*, ed. R. B. Braithwaite. London: Routledge and Kegan Paul.
—— 'Facts and Propositions', *Proceedings of the Aristotelian Society*, Supp. Vol. 7, pp. 153-70.

Russell, B. (1910), 'On the Nature of Truth and Falsehood', reprinted in *The Collected Papers of Bertrand Russell*, vol. 6, ed. J. G. Slater. London: Routledge, 1992.
—— (1922), 'Introduction' to Wittgenstein (1922).

Wittgenstein, L. (1913), *Notes on Logic*, in Wittgenstein (1979b).
—— (1922), *Tractatus Logico-Philosophicus*, trans. C.K.  Ogden. London: Routledge and Kegan Paul (second edition 1933).
—— (1979a), *Wittgenstein's Lectures, Cambridge 1932-1935*, ed. A. Ambrose.  Oxford: Blackwell.
—— (1979b), *Notebooks 1914-1916*, trans. G. E. M. Anscombe. Oxford: Blackwell.

# BIOGRAPHY

John Langshaw Austin (26 March 1911 – 8 February 1960) was a British philosopher of language. He is remembered primarily as the developer of the theory of speech acts. He read *Literae Humaniores* at Oxford and graduated with a first class honours degree in 1933. After serving in MI6 during World War II, Austin returned to Oxford as White's Professor of Moral Philosophy at Oxford, a post he retained up until his death in 1960. J.L. Austin was president of the Aristotelian Society from 1956 to 1957.

"Truth" was originally published in the *Proceedings of the Aristotelian Society, Supplementary Volume XXIV* (1950).

Charles Travis received a BA in philosophy from Berkeley and a PhD from UCLA. He has taught in four or five countries (to take no stands on how to count), The US, Canada, The Netherlands, Scotland and England), and many more universities. His last three positions have been at the University of Stirling, Northwestern University and King's College London. He has had visiting appointments at, most notably, the University of Michigan and Harvard, and, recently, as a visiting researcher at the University of Sydeney. He has also lectured in the Collège de France. He now lives in Portugal (Lavra) where he (mostly) writes, but is also co-organiser of a project, "The Bounds of Judgement", of the Mind Language and Action research group of the Institute of Philosophy of the University of Porto. In the beginning he wrote on philosophy of language, and is liable to do so again. At present he works primarily on Frege, Wittgenstein, philosophy of perception and philosophy of psychology. He is currently working on a book on Frege, to be titled: *Frege: The Invisible Realm*. He has published two books on Wittgenstein, one on propositional attitudes, and three collections of essays, the most recent titled, *Perception: Essays After Frege*.

# *Truth*

# J.L. AUSTIN

# TRUTH

## J.L. AUSTIN

1. "WHAT is truth?" said jesting Pilate, and would not stay for an answer. Pilate was in advance of his time. For "truth" itself is an abstract noun, a camel, that is, of a logical construction, which cannot get past the eye even of a grammarian. We approach it cap and categories in hand: we ask ourselves whether Truth is a substance (the Truth, the Body of Knowledge), or a quality (something like the colour red, inhering in truths), or a relation ("correspondence ")[1]. But philosophers should take something more nearly their own size to strain at. What needs discussing rather is the use, or certain uses, of the word "true." In *vino*, possibly, "*veritas*," but in a sober symposium "*verum*."

2. What is it that we say is true or is false? Or, how does the phrase "is true" occur in English sentences? The answers appear at first multifarious. We say (or are said to say) that beliefs are true, that descriptions or accounts are true, that propositions or assertions or statements are true, and that words or sentences are true: and this is to mention only a selection of the more obvious candidates. Again, we say (or are said to say) "It is true that the cat is on the mat," or "It is true to say that the cat is on the mat," or "'The cat is on the mat' is true." We also remark on occasion, when someone else has said something, "Very true" or "That's true" or "True enough."

Most (though not all) of these expressions, and others besides, certainly do occur naturally enough. But it seems reasonable to ask whether there is not some use of "is true" that is primary, or some generic name for that which at bottom we are always saying "is true." Which, if any, of these expressions is to be taken *au pied de la lettre*? To answer this will not take us long, nor, perhaps, far: but in philosophy the foot of the letter is the foot of the ladder.

I suggest that the following are the primary forms of expression : –

It is true (to say) that the cat is on the mat.

_____

[1] It is sufficiently obvious that "truth" is a substantive, "true" an adjective and "of" in "true of" a preposition.

That statement (of his, etc.) is true.

The statement that the cat is on the mat is true.

But first for the rival candidates.

*(a)* Some say that "truth is primarily a property of beliefs." But it may be doubted whether the expression "a true belief" is at all common outside philosophy and theology: and it seems clear that a man is said to hold a true belief when and in the sense that he believes (in) *something which is true*, or believes that *something which* is true is true. Moreover if, as some also say, a belief is "of the nature of a picture," then it is of the nature of what cannot be true, though it may be, for example, faithful.[2]

*(b)* True descriptions and true accounts are simply varieties of true statements or of collections of true statements, as are true answers and the like. The same applies to propositions too, in so far as they are genuinely said to be true (and not, as more commonly, sound, tenable and so on).[3]  A proposition in law or in geometry is something portentous, usually a generalisation, that we are invited to accept and that has to be recommended by argument: it cannot be a direct report on current observation – if you look and inform me that the cat is on the mat, that is not a proposition though it is a statement. In philosophy, indeed, "proposition" is sometimes used in a special way for "the meaning or sense of a sentence or family of sentences" : but whether we think a lot or little of this usage, a proposition in this sense cannot, at any rate, be what we say is true or false. For we never say "The meaning (or sense) of this sentence (or of these words) is true": what we do say is what the judge or jury says, namely that "*The words* taken in this sense, or if we assign to them such and such a meaning, or so interpreted or understood, *are true*."

(c) Words and sentences are indeed said to be true, the former often, the latter rarely. But only in certain senses. Words as discussed by philologists, or by lexicographers, grammarians, linguists, phoneticians, printers, critics (stylistic or textual) and so on, are not true or false: they

_____

[2] A likeness is true to life, but not true of it. A word picture can be true, just because it is not a picture.
[3] Predicates applicable also to "arguments," which we likewise do not say are true, but, for example, valid.

are wrongly formed, or ambiguous or defective or untranslatable or unpronouncable or misspelled or archaistic or corrupt or what not.[4] Sentences in similar contexts are elliptic or involved or alliterative or ungrammatical. We may, however, genuinely say "His closing words were very true" or "The third sentence on page 5 of his speech is quite false": but here "words" and "sentence" refer, as is shown by the demonstratives (possessive pronouns, temporal verbs, definite descriptions, etc.), which in this usage consistently accompany them, to the words or sentence *as used by a certain person on a certain occasion.* That is, they refer (as does "Many a true word spoken in jest") to *statements.*

A statement is made and its making is a historic event, the utterance by a certain speaker or writer of certain words (a sentence) to an audience with reference to a historic situation, event or what not.[5]

A sentence is made *up* of words, a statement is made *in* words. A sentence is not English or not good English, a statement is not in English or not in good English. Statements are made, words or sentences are used. We talk of *my* statement, but of *the English* sentence (if a sentence is mine, I coined it, but I don't coin statements). The *same* sentence is used in making *different* statements (I say "It is mine," you say "It is mine"): it may also be used on two occasions or by two persons in making the same statement, but for this the utterance must be made with reference to the same situation or event.[6] We speak of "the statement that S," but of "the sentence 'S'", not of "the sentence that S."[7]

_____

[4] Peirce made a beginning by pointing out that there are two (or three) different senses of the word "word," and adumbrated a technique ("counting" words) for deciding what is a "different sense." But his two senses are not well defined, and there are many more, – the "vocable" sense, the philologist's sense in which "grammar" is the same word as "glamour," the textual critic's sense in which the "the" in 1.254 has been written twice, and so on. With all his 65 divisions of signs, Peirce does not, I believe, distinguish between a sentence and a statement.

[5] "Historic" does not, of course, mean that we cannot speak of future or possible statements. A "certain" speaker need not be any definite speaker. "Utterance" need not be public utterance – the audience may be the speaker himself.

[6] "The same" does not always mean the same. In fact it has no meaning in the way that an "ordinary" word like "red" or "horse" has a meaning: it is a (the typical) device for establishing and distinguishing the meanings of ordinary words. Like "real," it is part of our apparatus *in* words for fixing and adjusting the semantics *of* words.

[7] Inverted commas show that the words, though uttered (in writing), are not to be taken as a statement by the utterer. This covers two possible cases, (i) where what is to be discussed is the sentence (ii) where what is to be discussed is a statement made elsewhere in the words "quoted." Only in case (i) is it correct to say simply that the token is doing duty for the type (and even here it is quite incorrect to say that "The cat

When I say that a statement is what is true, I have no wish to become wedded to one word. "Assertion," for example, will in most contexts do just as well, though perhaps it is slightly wider. Both words share the weakness of being rather solemn (much more so than the more general "what you said" or "your words"), – though perhaps we are generally being a little solemn when we discuss the truth of anything. Both have the merit of clearly referring to the historic use of a sentence by an utterer, and of being therefore precisely not equivalent to "sentence." For it is a fashionable mistake to take as primary "(The sentence) 'S' is true (in the English language)." Here the addition of the words "in the English language" serves to emphasize that "sentence" is not being used as equivalent to "statement," so that it precisely is not what can be true or false (and moreover, "true in the English language" is a solecism, mismodelled presumably, and with deplorable effect, on expressions like "true in geometry").

3. When is a statement true? The temptation is to answer (at least if we confine ourselves to "straightforward" statements): "When it corresponds to the facts." And as a piece of standard English this can hardly be wrong. Indeed, I must confess I do not really think it is wrong at all: the theory of truth is a series of truisms. Still, it can at least be misleading.

If there is to be communication of the sort that we achieve by language at all, there must be a stock of symbols of some kind which a communicator ("the speaker") can produce "at will " and which a communicatee ("the audience") can observe: these may be called the "words," though, of course, they need not be anything very like what we should normally call words – they might be signal flags, etc. There must also be something other than the words, which the words are to be used to communicate about: this may be called the "world." There is no reason why the world should not include the words, in every sense except the sense of the actual statement itself which on any particular occasion is being made about the world. Further, the world must exhibit (we must observe) similarities and dissimilarities (there could not be the one without the other): if everything were either absolutely indistinguishable from anything else or completely unlike anything else,

---

is on the mat" is the *name* of an English sentence, – though possibly *The Cat is on the Mat* might be the title of a novel, or a bull might be known as *Catta est in matta*). Only in case (ii) is there something true or false, *viz.* (not the quotation but) the statement made in the words quoted.

there would be nothing to say. And finally (for present purposes-of course there are other conditions to be satisfied too) there must be two sets of conventions: –

> *Descriptive* conventions correlating the words (= sentences) with the types of situation, thing, event, etc., to be found in the world.

> *Demonstrative* conventions correlating the words (= statements) with the historic situations, etc., to be found in the world.[8]

A statement is said to be true when the historic state of affairs to which it is correlated by the demonstrative conventions (the one to which it "refers") is of a type[9] with which the sentence used in making it is correlated by the descriptive conventions.[10]

3a. Troubles arise from the use of the word "facts" for the historic situations, events, etc., and in general, for the world. For "fact" is regularly used in conjunction with "that" in the sentences "The fact is that S" or "It is a fact that S" and in the expression "the fact that S," all of which imply that it would be true to say that S.[11]

_____

[8] Both sets of conventions may be included together under "semantics." But they differ greatly.

[9] "Is of a type with which" means "is sufficiently like those standard states of affairs with which." Thus, for a statement to be true one state of affairs must be like certain others, which is a natural relation, but also sufficiently like to merit the same "description," which is no longer a purely natural relation. To say "This is red" is not the same as to say "This is like those", nor even as to say "This is like those which were called red". That things are similar, or even "exactly" similar, I may literally see, but that they are the same I cannot literally see – in calling them the same colour a convention is involved additional to the conventional choice of the name to be given to the colour which they are said to be.

[10] The trouble is that sentences contain words or verbal devices to serve both descriptive and demonstrative purposes (not to mention other purposes), often both at once. In philosophy we mistake the descriptive for the demonstrative (theory of universals) or the demonstrative for the descriptive (theory of monads). A sentence as normally distinguished from a mere word or phrase is characterised by its containing a minimum of verbal demonstrative devices (Aristotle's "reference to time"); but many demonstrative conventions are non-verbal (pointing, etc.), and using these we can make a statement in a single word which is not a "sentence". Thus, "languages" like that of (traffic, etc.) *signs* use quite distinct media for their descriptive and demonstrative elements (the sign on the post, the site of the post). And however many verbal demonstrative devices we use as auxiliaries, there must *always* be a non-verbal *origin* for these coordinates, which is the point of utterance of the statement.

[11] I use the following *abbreviations* : –
    S    *for* the cat is on the mat.
    ST   *for* it is true that the cat is on the mat.
    tst  *for* the statement that.

This may lead us to suppose that

> (i) "fact" is only an alternative expression for "true statement." We note that when a detective says "Let's look at the facts" he doesn't crawl round the carpet, but proceeds to utter a string of statements we even talk of "stating the facts";

> (ii) for every true statement there exists "one" and its own precisely corresponding fact – for every cap the head it fits.

It is (i) which leads to some of the mistakes in "coherence" or formalist theories; (ii) to some of those in "correspondence" theories. Either we suppose that there is nothing there but the true statement itself, nothing to which it corresponds, or else we populate the world with linguistic *Doppelgänger* (and grossly overpopulate it – every nugget of "positive" fact overlaid by a massive concentration of "negative" facts, every tiny detailed fact larded with generous general facts, and so on).

When a statement is true, there is, *of course*, a state of affairs which makes it true and which is *toto mundo* distinct from the true statement about it: but equally of course, we can only describe that state of affairs in *words* (either the same or, with luck, others). I can only describe the situation in which it is true to say that I am feeling sick by saying that it is one in which I am feeling sick (or experiencing sensations of nausea)[12]: yet between stating, however, truly that I am feeling sick and feeling sick there is a great gulf fixed.[13]

"Fact that" is a phrase designed for use in situations where the distinction between a true statement and the state of affairs about which it is a truth is neglected; as it often is with advantage in ordinary life,

---

I take tstS as my example throughout and not, say, tst Julius Caesar was bald or tst all mules are sterile, because these latter are apt in their different ways to make us overlook the distinction between sentence and statement: we have, apparently, in the one case a sentence capable of being used to refer to only one historic situation, in the other a statement without reference to at least (or to any particular) one.

If space permitted other types of statement (existential, general, hypothetical, etc.) should be dealt with: these raise problems rather of meaning than of truth, though I feel uneasiness about hypotheticals.

[12] If this is what was meant by "'It is raining' is true if and only if it is raining," so far so good.

[13] It takes two to make a truth. Hence (obviously) there can be no criterion of truth in the sense of some feature detectable in the statement itself which will reveal whether it is true or false. Hence, too, a statement cannot without absurdity refer to itself.

though seldom in philosophy – above all in discussing truth, where it is precisely our business to prise the words off the world and keep them off it. To ask "Is the fact that S the true statement that S or that which it is true of?" may beget absurd answers. To take an analogy: although we may sensibly ask "Do we *ride* the word 'elephant' or the animal?" and equally sensibly "Do we *write* the word or the animal?" it is nonsense to ask "Do we *define* the word or the animal?" For defining an elephant (supposing we ever do this) is a compendious description of an operation involving both word and animal (do we focus the image or the battleship?); and so speaking about "the fact that" is a compendious way of speaking about a situation involving both words and world.[14]

3b. "Corresponds" also gives trouble, because it is commonly given too restricted or too colourful a meaning, or one which in this context it cannot bear. The only essential point is this: that the correlation between the words (= sentences) and the type of situation, event, etc., which is to be such that when a statement in those words is made with reference to a historic situation of that type the statement is then true, is *absolutely and purely* conventional. We are absolutely free to appoint *any* symbol to describe *any* type of situation, so far as merely being true goes. In a small one-spade language tst nuts might be true in exactly the same circumstances as the statement in English that the National Liberals are the people's choice.[15] 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; a statement no more needs, in order to be true, to reproduce the "multiplicity," say, or the "structure" or "form" of the reality, than a word needs to be echoic or writing pictographic. To suppose that it does, is to fall once again into the error of reading back into the world the features of language.

The more rudimentary a language, the more, very often, it will tend to have a "single" word for a highly "complex" type of situation: this has such disadvantages as that the language becomes elaborate to learn and is incapable of dealing with situations which are non-standard, unforeseen, for which there may just be no word. When we go abroad

_____

[14] "It is true that S" and "It is a fact that S" are applicable in the same circumstances; the cap fits when there is a head it fits. Other words can fill the same role as "fact"; we say, *e.g.,* "The situation is that S."

[15] We could use "nuts" even now as a code-word: but a code, as a transformation of a language, is distinguished from a language, and a code-word despatched is not (called) "true".

equipped only with a phrase-book, we may spend long hours learning by heart –

A$^l$-moest-fa$^i$nd-$^e$tschâ$^r$woum$^e$n,
Ma$^l$-hwîl-iz-wau$^r$pt (bènt),

and so on and so on, yet faced with the situation where we have the pen of our aunt, find ourselves quite unable to say so. The characteristics of a more developed language (articulation, morphology, syntax, abstractions, etc.), do not make statements in it any more capable of being true or capable of being any more true, they make it more adaptable, more learnable, more comprehensive, more precise and so on; and *these* aims may no doubt be furthered by making the language (allowance made for the nature of the medium) "mirror" in conventional ways features descried in the world.

Yet even when a language does "mirror" such features very closely (and does it ever?) the truth of statements remains still a matter, as it was with the most rudimentary languages, of the words used being the ones *conventionally appointed* for situations of the type to which that referred to belongs. A picture, a copy, a replica, a photograph – these are *never* true in so far as they are reproductions, produced by natural or mechanical means: a reproduction can be accurate or lifelike (true *to* the original), as a gramophone recording or a transcription may be, but not true (*of*) as a record of proceedings can be. In the same way a (natural) sign *of* something can be infallible or unreliable but only an (artificial) sign *for* something can be right or wrong.[16]

There are many intermediate cases between a true account and a faithful picture, as here somewhat forcibly contrasted, and it is from the study of these (a lengthy matter) that we can get the clearest insight into the contrast. For example, maps: these may be called pictures, yet they are highly conventionalised pictures. If a map can be clear or accurate or misleading, like a statement, why can it not be true or exaggerated? How do the "symbols" used in map-making differ from those used in statement-making? On the other hand, if an air-mosaic is not a map, why is it not? And when does a map become a diagram? These are the really illuminating questions.

_____

[16] Berkeley confuses these two. There will not be books in the running brooks until the dawn of hydro-semantics.

4. Some have said that –

> To say that an assertion is true is not to make any further assertion at all.

> In all sentences of the form "p is true" the phrase "is true" is logically superfluous.

> To say that a proposition is true is just to assert it, and to say that it is false is just to assert its contradictory.

But wrongly. TstS (except in paradoxical cases of forced and dubious manufacture) refers to the world or any part, of it exclusive of tstS, *i.e.*, of itself.[17] TstST refers to the world or any part of it *inclusive* of tstS, though once again exclusive of itself, *i.e.*, of tstST. That is, tstST refers to something to which tstS cannot refer. TstST does not, certainly, include any statement referring to the world exclusive of tstS which is not included already in tstS – more, it seems doubtful whether it does include that statement about the world exclusive of tstS which is made when we state that S. (If I state that tstS is true, should we really agree that I have stated that S? Only "by implication."[18]) But all this does not go any way to show that tstST is not a statement different from tstS. If Mr. Q writes on a noticeboard "Mr. W is a burglar," then a trial is held to decide whether Mr. Q's published statement that Mr. W is a burglar is a libel: finding "Mr. Q's statement was true (in substance and in fact)." Thereupon a second trial is held, to decide whether Mr. W is a burglar, in which Mr. Q's statement is no longer under consideration: verdict "Mr. W is a burglar." It is an arduous business to hold a second trial: why is it done if the verdict is the same as the previous finding?[19]

_____

[17] A statement may refer to "itself " in the sense, *e.g.*, of the sentence used or the utterance uttered in making it ("statement" is not exempt from all ambiguity). But paradox does result if a statement purports to refer to itself in a more full-blooded sense, purports, that is, to state that it itself is true, or to state what it itself refers to ("This statement is about Cato").

[18] And "by implication" tstST asserts something about the making of a statement which tstS certainly does not assert.

[19] This is not quite fair: there are many legal and personal reasons for holding two trials, – which, however, do not affect the point that the issue being tried is not the same.

What is felt is that the evidence considered in arriving at the one verdict is the same as that considered in arriving at the other. This is not strictly correct. It is more nearly correct that whenever tstS is true then tstST is also true and conversely, and that whenever tstS is false tstST is also false and conversely.[20] And it is argued that the words "is true" are logically superfluous because it is believed that generally if any two statements are always true together and always false together then they must mean the same. Now whether this is in general a sound view may be doubted: but even if it is, why should it not break down in the case of so obviously "peculiar" a phrase as "is true"? Mistakes in philosophy notoriously arise through thinking that what holds of "ordinary" words like "red" or "growls" must also hold of extraordinary words like "real" or "exists." But that "true" is just such another extraordinary word is obvious.[21]

There is something peculiar about the "fact" which is described by tstST, something which may make us hesitate to call it a "fact" at all; namely, that the relation between tstS and the world which tstST asserts to obtain is a *purely conventional* relation (one which "thinking makes so"). For we are aware that this relation is one which we could alter at will, whereas we like to restrict the word "fact" to *hard* facts, facts which are natural and unalterable, or anyhow not alterable at will. Thus, to take an analogous case, we may not like calling it a fact that the word elephant means what it does, though we can be induced to call it a (soft) fact – and though, of course, we have no hesitation in calling it a fact that contemporary English speakers use the word as they do.

An important point about this view is that it confuses falsity with negation: for according to it, it is the same thing to say "He is not at home" as to say "It is false that he is at home." (But what if no one has said that he *is* at home? What if he is lying upstairs dead?) Too many philosophers maintain, when anxious to explain away negation, that a negation is just a second order affirmation (to the effect that a certain first order affirmation is false), yet, when anxious to explain away falsity, maintain that to assert that a statement is false is just to assert its negation (contradictory). It is impossible to deal with so fundamental a

_____

[20] Not *quite* correct, because tstST is only in place at all when tstS is envisaged as made and has been verified.

[21] *Unum, verum, bonum*, – the old favourites deserve their celebrity. There is something odd about each of them. Theoretical theology is a form of onomatolatry.

matter here.[22] Let me assert the following merely. Affirmation and negation are exactly on a level, in this sense, that no language can exist which does not contain conventions for both and that both refer to the world equally directly, not to statements about the world: whereas a language can quite well exist without any device to do the work of "true" and "false." Any satisfactory theory of truth must be able to cope equally with falsity[23]: but "is false" can only be maintained to be logically superfluous by making this fundamental confusion.

5. There is another way of coming to see that the phrase "is true" is not logically superfluous, and to appreciate what sort of a statement it is to say that a certain statement is true. There are numerous other adjectives which are in the same class as "true" and "false," which are concerned, that is, with the relations between the words (as uttered with reference to a historic situation) and the world, and which nevertheless no one would dismiss as logically superfluous. We say, for example, that a certain statement is exaggerated or vague or bald, a description somewhat rough or misleading or not very good, an account rather general or too concise. In cases like these it is pointless to insist on deciding in simple terms whether the statement is "true or false." Is it true or false that Belfast is north of London? That the galaxy is the shape

_____

[22] The following two sets of logical axioms are, as Aristotle (though not his successors) makes them, quite distinct: –
(a) No statement can be both true and false.
No statement can be neither true nor false.
(b) Of two contradictory statements –
        Both cannot be true.
        Both cannot be false.
The second set demands a definition of contradictories, and is usually joined with an unconscious postulate that for every statement there is one and only one other statement such that the pair are contradictories. It is doubtful how far any language does or must contain contradictories, however defined, such as to satisfy both this postulate and the set of axioms (*b*).
Those of the so-called "logical paradoxes" (hardly a genuine class) which concern "true" and "false" are not to be reduced to cases of self-contradiction, any more than "S but I do not believe it" is. A statement to the effect that it is itself true is every bit as absurd as one to the effect that it is itself false. There are other types of sentence which offend against the fundamental conditions of all communication in ways distinct from the way in which "This is red and is not red" offends, – *e.g.*, "This does (I do) not exist," or equally absurd "This exists (I exist)." There are more deadly sins than one; nor does the way to salvation lie through any hierarchy.
[23] To be false is (not, of course, to correspond to a non-fact, but) to mis-correspond with a fact. Some have not seen how, then, since the statement which is false does not describe the fact with which it mis-corresponds (but misdescribes it), we know which fact to compare it with: this was because they thought of all linguistic conventions as descriptive, – but it is the demonstrative conventions which fix which situation it is to which the statement refers. No statement can state what it itself refers to.

of a fried egg? That Beethoven was a drunkard? That Wellington won the battle of Waterloo? There are various *degrees and dimensions* of success in making statements: the statements fit the facts always more or less loosely, in different ways on different occasions for different intents and purposes. What may score full marks in a general knowledge test may in other circumstances get a gamma. And even the most adroit of languages may fail to "work" in an abnormal situation or to cope, or cope reasonably simply, with novel discoveries: is it true or false that the dog goes round the cow?[24] What, moreover, of the large class of cases where a statement is not so much false (or true) as out of place, *inept* ("All the signs of bread" said when the bread is before us)?

We become obsessed with "truth" when discussing statements, just as we become obsessed with "freedom" when discussing conduct. So long as we think that what has always and alone to be decided is whether a certain action was done freely or was not, we get nowhere: but so soon as we turn instead to the numerous other adverbs used in the same connexion ("accidentally," "unwillingly," "inadvertently," etc.), things become easier, and we come to see that no concluding inference of the form "Ergo, it was done freely (or not freely)" is required. Like freedom, truth is a bare minimum or an illusory ideal (the truth, the whole truth and nothing but the truth about, say, the battle of Waterloo or the *Primavera*).

6. Not merely is it jejune to suppose that all a statement aims to be is "true," but it may further be questioned whether every "statement" does aim to be true at all. The principle of Logic, that "Every proposition must be true or false," has too long operated as the simplest, most persuasive and most pervasive form of the descriptive fallacy. Philosophers under its influence have forcibly interpreted all "propositions" on the model of the statement that a certain thing is red, as made when the thing concerned is currently under observation.

_____

[24] Here there is much sense in "coherence" (and pragmatist) theories of truth, despite their failure to appreciate the trite but central point that truth is a matter of the relation between words and world, and despite their wrong-headed *Gleichschaltung* of all varieties of statemental failure under the one head of "partly true" (thereafter wrongly equated with "part of the truth"). "Correspondence" theorists too often talk as one would who held that every map is either accurate or inaccurate; that accuracy is a single and the sole virtue of a map; that every country can have but one accurate map; that a map on a larger scale or showing different features must be a map of a different country; and so on.

Recently, it has come to be realized that many utterances which have been taken to be statements (merely because they are not, on grounds of grammatical form, to be classed as commands, questions, etc.) are not in fact descriptive, nor susceptible of being true or false. When is a statement not a statement? When it is a formula in a calculus: when it is a performatory utterance: when it is a value-judgment: when it is a definition: when it is part of a work of fiction – there are many such suggested answers. It is simply not the business of such utterances to "correspond to the facts" (and even genuine statements have other businesses besides that of so corresponding).

It is a matter for decision how far we should continue to call such masqueraders "statements" at all, and how widely we should be prepared to extend the uses of "true" and "false" in "different senses." My own feeling is that it is better, when once a masquerader has been unmasked, *not* to call it a statement and *not* to say it is true or false. In ordinary life we should not call most of them statements at all, though philosophers and grammarians may have come to do so (or rather, have lumped them all together under the term of art "proposition"). We make a difference between "You said you promised" and "You stated that you promised": the former can mean that you said "I promise," whereas the latter must mean that you said "I promised": the latter, which we say you "stated," is something which is true or false, whereas for the former, which is not true or false, we use the wider verb to "say." Similarly, there is a difference between "You say this is (call this) a good picture" and "You state that this is a good picture." Moreover, it was only so long as the real nature of arithmetical formulae, say, or of geometrical axioms remained unrecognised, and they were thought to record information about the world, that it was reasonable to call them "true" (and perhaps even "statements," – though were they ever so called?): but once their nature has been recognized, we no longer feel tempted to call them "true" or to dispute about their truth or falsity.

In the cases so far considered the model "This is red" breaks down because the "statements" assimilated to it are not of a nature to correspond to facts at all, – the words are not descriptive words, and so on. But there is also another type of case where the words *are* descriptive words and the "proposition" does in a way have to correspond to facts, but precisely not in the way that "This is red" and similar statements setting up to be true have to do.

In the human predicament, for use in which our language is designed, we may wish to speak about states of affairs which have not been observed or are not currently under observation (the future, for example). And although we *can* state anything "as a fact" (which statement will then be true or false[25]) we need not do so: we need only say "The cat *may be* on the mat." This utterance is quite different from tstS, – it is not a statement at all (it is not true or false; it is compatible with "The cat may *not* be on the mat"). In the same way, the situation in which we discuss whether and state that tstS is *true* is different from the situation in which we discuss whether it is *probable* that S. Tst it is probable that S is out of place, inept, in the situation where we can make tstST, and, I think, conversely. It is not our business here to discuss probability: but is worth observing that the phrases "It is true that" and "It is probable that" are in the same line of business,[26] and in so far incompatibles.

7. In a recent article in *Analysis* Mr. Strawson has propounded a view of truth which it will be clear I do not accept. He rejects the "semantic" account of truth on the perfectly correct ground that the phrase "is true" is not used in talking about *sentences*, supporting this with an ingenious hypothesis as to how meaning may have come to be confused with truth: but this will not suffice to show what he wants, – that "is true" is not used in talking about (or that "truth is not a property of") *anything*. For it is used in talking about *statements* (which in his article he does not distinguish clearly from sentences). Further, he supports the "logical superfluity" view to this extent, that he agrees that to say that ST is not to make any further assertion at all, beyond the assertion that S: but he disagrees with it in so far as he thinks that to say that ST *is* to *do* something more than just to assert that S, – it is namely to *confirm* or to *grant* (or something of that kind) the assertion, made or taken as made already, that S. It will be clear that and why I do not accept the first part of this: but what of the second part? I agree that to say that ST "is" very often, and according to the all-important linguistic occasion, to confirm tstS or to grant it or what not; but this cannot show that to say that ST is not also and at the same time to make an assertion about tstS. To say that I believe you "is" on occasion to accept your statement; but it is also to make an assertion, which is not made by the

_____

[25] Though it is not yet in place to call it either.  For the same reason, one cannot lie or tell the truth about the future.
[26] Compare the odd behaviours of "was" and "will be" when attached to "true" and to "probable."

strictly performatory utterance "I accept your statement." It is common for quite ordinary statements to have a performatory "aspect": to say that you are a cuckold may be to insult you, but it is also and at the same time to make a statement which is true or false. Mr. Strawson, moreover, seems to confine himself to the case where I say "Your statement is true" or something similar, – but what of the case where you state that S and I *say* nothing but "*look and see*" that your statement is true? I do not see how this critical case, to which nothing analogous occurs with strictly performatory utterances, could be made to respond to Mr. Strawson's treatment.

One final point: if it is admitted (*if*) that the rather boring yet satisfactory relation between words and world which has here been discussed does genuinely occur, why should the phrase "is true" not be our way of describing it? And if it is not, what else is?

*As a Matter of Fact*

CHARLES TRAVIS

# AS A MATTER OF FACT

## CHARLES TRAVIS

THE word 'true', Frege tells us, is not a relation word. (1918: 59) Nor, he also tells us, is it *quite* right to call truth a property (though he will do so *pro tem*). Truth is not a relation. Fine. Though relations can degenerate. Being older than Methusaleh (if as reputed) is a property none of us has. If Methusaleh is history, then to have it would be to relate in a certain way to Methusaleh. No one else one might relate to would do the trick. Similarly, if to be true were to relate to something (what the truth-bearer was true of), what Frege's point suggests is that there could be only *one* thing for this relatum to be, no matter what the truth-bearer. And indeed, construing truth as a relation would leave only one such eligible candidate. Truth, in any case, comes on the scene along with a certain relation: that of (a representer) representing something as something. It is such representing which is done truly or falsely, the representing thus done accordingly true or false.There is truth, one might think, just where the third term in this relation—the way things were represented as being—related suitably to the second term—what was so represented. Where there is truth outright, there is but *one* thing for this second term to be. It is the way things are which is represented, truly or falsely, as (things) being such-and-such way. It is thus relating suitably (or as required) to *this* (the way things are) that makes for truth wherever there is truth outright. Always the same relatum, just as it is always the same relatum in being older than Methusaleh. One might give this relatum different names. Perhaps 'things' would do, construing 'things' catholicly enough. 'Things being as they are' might be more suggestive. One might also speak here of 'the world', or 'history'. In any case, one might enquire as to what *relating suitably* would be here. Such is Austin's question. Nothing in Frege's suggestion rules it out.

The role of what Frege calls a thought is to be, as he puts it, "that by which truth can come into question at all." (1918: 60) If representing is an act, or, as in representing to oneself, a stance, well, thoughts can neither act nor hold stances. But there is an aspect of the verb on which Frege's *Gedanke* might serve as a first term in the relation, representing-as. On this aspect, it is enough for expressing it (assertively) to be representing truly. By this avenue it lets itself in for truth or falsity. Unlike agents (or content-bearers), for it to be the one it is is for it to represent what it does as the way it does. A thought can so serve, though, only if, in this aspect, it stops nowhere short of representing-as. It thus contrasts with a concept, which *does* stop short. The concept *nonchalant* may well be of a way Sid is. We might thus say that it is true of him. But Sid is none of its business. It does not depend for its

existence on there so much as being Sid. It neither represents him as nonchalant or not. My example is a one-place concept. But the point would hold for any n. It would hold for n=0 if we chose to recognise zero-place concepts. (Such a concept would, e.g., be of things being such that Sid is nonchalant—once again that catholic 'things'.) In sum: for a thought to be true is (harmlessly) for it to be true of the way things are. There is *something* relational in that.

Frege argues against a correspondence theory of truth—something Austin will have no truck with. On such an ill-begotten theory, there is a domain of truth-bearers (thoughts, or what plays their role); and there is a distinct domain of multitudinous items of some other sort. There is then a relation between these domains such that a truth-bearer is true just in case there is an item in this second domain to which it so relates. Such a theory gets grammar all wrong at the very start. We will soon see why Austin *could have* no truck with it.

Frege tells us that the content of the word 'true' is unique and undefinable. (1918: 60) not that 'true' has no, or no identifiable, content. He tells us, for a start, that its content is unfolded (in most general respects) by the laws of truth (that is, of logic). (1897: 139) At the same time, he also asserts that truth is an identity under predication: predicate truth of a thought, and you get that thought back. So, if 'is true' *adds* nothing, perhaps after all it has no content? Or is the point rather that looking at predication is looking in the wrong place to find that content? Here is Frege on this:

> Thus it is to be observed that the relation of thoughts to truth is not to be compared to that of subject to predicate. Subject and predicate are, to be sure, thought-elements (understood in the logical sense); they stand on the same level with respect to knowledge. By putting together subject and predicate one always arrives only at a thought, never from a *Sinn* to its *Bedeutung*, never from a thought to its truth value. One moves around on the same level, but never steps from the one level to the next. A truth value cannot be a thought-element, just as little as, say, the sun, because it is not a *Sinn*, but rather an object. (1892: 34–35)

Representing something *as* something is one thing. To represent something *to be* something is to take a further step; one which need not be taken merely to take the first. If Sid were suave, he might be rich. Such is not to suppose he *is* suave. Part of Frege's point is: one can never get from representing-as to representing-to-be merely by engaging in more representing-as. Add all the representing-as you like to that with which you started, and the result is only more of it. The step to representing-to-be remains untaken. As Frege puts this, one never thus takes 'the step from *Sinn* to *Bedeutung*; from a truth-bearer to its truth-value.'

To step from *Sinn* to *Bedeutung* is, where the *Sinn* is a thought, to commit, or acknowledge (*anerkennen*) the thought's credentials; its faithfulness to the way things are. This is something one does, if at all, only under a certain kind of compulsion: he must see there as nothing else for one suitably informed (thus for one in his position) to think. What is *felt* is rational compulsion. The only thing to think *can* be read: the only thing to think in pursuit of the goal truth. Frege (1897 loc cit) portrays the laws of logic as a *partial* answer to the question 'How must I think to reach the goal truth?'. One aims to take the step from *Sinn* to *Bedeutung* as directed by the answer to that question—not just the *partial* answer logic gives, but the full thing. Logic concerns itself with relations between ways for things to be represented as being. When one takes Frege's step, e.g., *in re* the thought that Sid smokes, his interest is in the whole story, but, most pressingly, the part logic (of necessity) leaves undiscussed: how pursuit of the goal truth is to go where it is a question of relating that which is represented as some way or other to ways for it to be represented—relating, that is, things (catholic reading) to ways to represent them. The difference between Frege and Austin on truth, viewed one way, parallels that between these two sorts of concerns.

Laws of logic concern exclusively items distinguished by a certain sort of generality. For each there is something it would be (for an item) to be, not identical with, but a case of it (not that way for things to be, such that Sid smokes, but a case of something being such as to smoke). A case—something which might be represented as being some way—has no such generality. Nor, accordingly, does it entail or probablify anything. Only its instancing one thing or another could do that. There had better be such a thing as what it would be to proceed so as to reach the goal truth in matters of what is a case of what—of just when a given generality is to be counted as instanced. Otherwise there is never such a thing as the (worldly) thing to think; truth is abolished. There must, that is, be such a thing as when that Sid smokes Murads would be the thing to judge in pursuing truth. Austin's interest is in what there is to say as to what pursuit of truth would be, particularly in connecting the above two terms of the representing-as relation.

What Austin has to say about this is found primarily (but not exclusively) in two places: the essay, "Truth" (1950) and his treatise on the subject, *How To Do Things With Words*. (1962 (*Nachlass*).) In the treatise he announces his intention to 'play Old Harry' with two traditional distinctions: true/false, and fact/value. The Old Harry he plays is the upshot of a failure to draw another dichotomy: a supposed one between two sorts of acts of representing-as: 'constatives' (acts of representing-to-be) and 'performatives' (bringing something about in, or by, saying something). Here is the conclusion Austin draws:

> The doctrine of the performative/constative distinction stands to the doctrine of locutionary and illocutionary acts in the total speech act as the *special* theory to the *general* theory. And the need for the general theory arises simply because the traditional 'statement' is an abstraction, an ideal, and so is its traditional truth or falsity. ...
>
> Stating, describing, etc., ... have no unique position over the matter of being related to facts in a unique way called being true or false, because truth and falsity are (except by an artificial abstraction which is always possible and legitimate for certain purposes) not names for relations, qualities, or what not, but for a dimension of assessment—how the words stand in respect of satisfactoriness to the facts, events, situations, etc., to which they refer. (1962: 147–8)
>
> It is essential to realise that 'true' and 'false', like 'free' and unfree', do not stand for anything simple at all; but only for a general dimension of being a right or proper things to say as opposed to a wrong thing, in these circumstances, to this audience, for these purposes and with these intentions. (op. cit.: 144)
>
> By the same token, the familiar contrast of 'normative or evaluative' as opposed to the factual is in need, like so many dichotomies, of elimination. (op. cit.: 148)

What we are dealing with here, Austin concludes, is a continuum; most notably a continuum in terms and standards of evaluation of acts of representing as successes or failures, and of the terms in which particular ones are to be evaluated. Notions with some evaluative core, such as fair or just, or close relatives, show up, for example, in questions as to whether something was a fair description of how things were, or 'true to the facts', or, in the circumstances of the act, would give a just impression, or well serve the purposes to which the act might be expected to be put. Should the description, 'The street is lined with eating establishments', e.g., given of a street full of soup kitchens interspersed with gin mills affording the odd free pickled egg, count as a just enough account of how things were to merit the title 'true'?—a question whose answer is more than likely to depend on the circumstances in which the description was, or would be, given.

Here we see how it is essential to the picture Austin tries to paint that correspondence theories of truth be non-starters. For, as such theories conceive things, there are two distinct autonomous domains, each of whose denizens are what they are independent of any such considerations as what it would be fair, or just to say, or what would mislead, or what might be a better or worse description of the facts. Truth is then merely a matter of whether, for an item in the first domain, there is an item in the second which is a match. There is no room here for evaluations to depend on Austinian considerations.

Frege focussed on the demands of logic: the demands it imposes on pursuit of truth, hence on thinking altogether; so, too, the demands imposed on there being anything for truth to impose demands on. On this last topic he wrote,

> A concept that is not sharply defined is wrongly termed a concept. Such quasi-conceptual constructions cannot be recognized as concepts by logic; it is impossible to lay down precise laws for them. The law of excluded middle is really just another form of the requirement that the concept should have a sharp boundary. Any object Δ that you choose to take either falls under the concept Φ or does not fall under it, *tertium non datur*. (Frege, 1903, §56)

The point applies to n-place concepts for any n, hence to zero-place concepts, hence to thoughts. Thoughts to which logic applied would then be ones tailored so as to respect it. Does such rule out room in the notion of truth for Austinian considerations?

Frege introduces a thought as 'that by which truth can come into question at all.' The 'that' here should be read as 'precisely that'—no more nor less. A thought, so to speak, is, or fixes, a *pure* question of truth. (The trouble with judgeable content—the notion *thought* replaced—is precisely that it failed this condition, since more (e.g., a truth value) was involved in being one of these than worked to fix a question of truth.) A further idea: if a thought has done this, then the answer to that question ('True of false?') can depend on *nothing* extra to the thought itself except that which the thought represents *as* something—things (catholically), the world. This *would* leave no room for Austinian considerations. Or at least none in matters of the truth of a *thought*, once that thought is identified properly as the one it is.

But perhaps there are other places to look? Perhaps the notion of truth is already involved, essentially, in there being any question of a thought having been expressed, or of it being one thought rather than another which was expressed on some occasion. Such an idea would fit well with what Austin has to say. Austin's starting point differs from Frege's. His focus is on historical acts of representing rather than on thoughts. This suggests a development for the above idea. First an observation. *Expressing* thoughts differs from having, or thinking, them: two different forms of representing-as. Expressing a thought is (an act of) making representing recognisable. A plausible thought: for *such* representing to be *is* for it to be recognisable; for it to be the representing it is is for it to be recognisable as that. Recognisable by whom? By those competent enough and suitably placed to do so. In the case of *our* representing, presumably by *us*, if suitably placed, and suitably *au fait* with the relevant ways for things to be (those in terms of which things were represented as they were). No *analysis* is on offer

here. But let us try to use what we know.

To engage in representing-as is to represent something as some way there is for things to be; thus as something with a certain sort of generality—something instanced (or counter-instanced) by a range of cases (some determinate range, one might hope). To grasp what way things were represented as being is (*inter alia*, perhaps) to grasp to what range of cases it would reach—when there would be a case of things being that way. Suppose, then, that Pia, lamenting Sid's love of lager, remarks to Zoë, 'Sid waddles'. Since for this form of representing-as to be is for it to be made recognisable, we must ask what *was* made recognisable in Pia's speaking as she did?

To answer this, we might turn first to the words she used. Those of us who know (enough) English will recognise these as speaking of the one named (if such there is) as a waddler. Of those of us who thus know what way for a thing to be is (so far) in question one may ask: what do we thus know as to to what cases this way reaches—as to what would, what would not, be a case of a thing so being? There *are* things we do know—perhaps for a start, say, that *penguins* waddle; so that if you choose a normal enough penguin, *there* will be a case of a waddler. Austin argues, though, that inevitably we will come up against cases where all there is (for the knowledgeable) to say—the *right* thing to say—can only be: 'Well, you *could* call that being a waddler. Or you could refuse to. Either would be compatible with all there is to know as to what being a waddler is.' If, in this domain, to be is to be recognisable, then so far there is all this to say, and no more. So far, Pia represented things as a way which some things would be cases of, some things would not; and, as for the rest, neither the answer 'Yes', nor the answer 'No', is mandated by the facts.

Such is stage one of our development. For stage two, a further observation. Just as with every human birth a new multitude of thoughts come into existence—thoughts of that new human that *he* is thus and so—so with every act of representing-as, a new way for things to be comes into existence: being as thus represented. So it is with Pia's act, of which we can now ask whether there is any more to say as to what it would be to be this new way for things to be—being as *she* represented them—than has been said already in discussing the words she used and what they speak of—being a waddler. Austin's answer is that there may well be. If one *could*, say, call what Sid does waddling, or doing it as much as he does being a waddler, and one could refuse to do so, either compatible with those words meaning what they do, perhaps, using words in the particular way Pia did, Sid's comporting himself as he does *ought* to be called being a waddler. Or ought not to.

What should count as waddling as *Pia* spoke of this? What would be a case? To answer this is to fix the demands on things being as she said; thus the standards of *truth* to which she is to be held. It is just here where Austinian considerations come into play. If Sid waddles when drunk, but only as his variant to putting a lampshade on his head, is it really fair to the facts to describe him as a waddler—say, in a discussion of the dire effects of alcohol's hidden calories? If Sid ceases to waddle when encased in sufficient 'supportive' garments, is it really fair to hold Pia responsible, in describing him as she did, for things being otherwise? If, for the sort of representing Pia's was—representing-as in *expressing* thoughts—to be is to be recognisable, then here is the arena in which Austin's points need, and seem, to hold good.

To sum up, Austin's concern is with that form of representing for which to be is to be recognisable. To render that concern in terms of Frege's notion of a thought (or the version of that tailored to meet Frege's demands on logic's applicability), one *could* say: it is a concern with what it would be for a given thought (in either of these senses) to have been expressed (or not) in a given such act of representing. The thought is to abstract from the act just that which determines when it would be (or have been) a case of representing *truly*. Austin's question is to what standards what is abstracted from—the concrete act—is to be held accountable for this. One might also ask for what representing it is to be held accountable; just what representations it is to be held responsible for having made. Is it to be held committed to more than is so in things being as the are? If, e.g., in the act Sid was described as a smoker, is the act to be held to have committed to more than is so given the way things are, or, e.g., more than would be so if Sid smoked only at his club on Fridays, on the balcony, after dinner? The answers to such questions, Austin plausibly enough holds, depend on the kinds of considerations he gestures at: whether, in the circumstances, it would be *fair* to describe Sid as a smoker if this is all he does, or whether one would have had the right to suppose that *more* than this was so if what she said were so; whether, in the circumstances, it would be fair to hold the agent (Pia) to have committed to more than is so if things are thus; that is, to have incurred liability to any failure here suffered, where this is, more specifically, failure to be representing truly. Questions like this, the point is, *are* questions as to what is to count as true. It is just that answers to them are already presupposed when we come to talk in terms of thoughts, in either of the above senses, at all.

At which point Frege and Austin can be seen as, as to the facts, not fundamentally at odds. Frege and Austin are equally concerned with what truth is, and *thereby* with questions of the form 'How must I think to reach the goal truth?'. They are just concerned with different places in which such questions can arise: Frege with relations among items with that certain sort of generality discussed above: for Frege, ways to

represent things as being, and what so represents them; for Austin relations between *what* is so represented and cases of so representing it.

Frege once suggested that we could "understand by the existence of a thought that it can be grasped by different thinkers as the same." (1919:146). A thought, on this conception, is *just* that (anything) which can be agreed to or disputed, whose truth can be investigated or wondered over, by many. A thought so conceived is identified by what is thus of interest to us—e.g., whether penguins mate in the spring. There is thus a thought identifiable as the thought that they do. The generality of such a thought is just that which relevant agreement can identify. It is then a *substantial* question whether such a thought fixes a *pure* question of truth on either of the notions of purity scouted above. *Perhaps* Frege and Austin disagree on the answer to this last question. Whether this is so or not, the concerns of each with truth are recognisable as legitimate, and in each case pursuable in the way each undertakes, provided one sees correctly how those very different ways connect with each other.

All of the above can be summed up as follows. A thought, one idea is, is *precisely* that by which truth comes into question at all, no more no less; so to speak, a *pure* question of truth. A further extension: Frege's step from *Sinn* to truth-value—from mere representing-as (as, e.g., in wondering) to full representing-to-be (as, e.g., in judging)—thus starts from something which leaves *nothing* undetermined as to when things would be as represented; it remaining only for the world, what *is* so represented, to speak. Holding fast to this conception, and looking for truth's content in its role in Frege's step, there *should* be little for us to find—certainly nothing like a role for truth in weighing up Austinian considerations. But if all this *defines* 'pure question of truth' whether there *are* any such becomes a substantial matter. It is now a *thesis* that such can be identified in speaking of, e.g., the thought that Sid smokes (or that penguins waddle)—or, for that matter, in any way which makes no reference to a concrete act of expressing the thought in question.

Such is one form of Austin's point. Holding fast to the above conception, the point can take another form. If there is a role for truth in Frege's step from *Sinn* to *Bedeutung,* then such is *one* place to look for truth's content. But if thoughts are thus abstracted from acts of representing, another place to look is in the abstracting. To know what thought Pia expressed in describing Sid as a smoker is to know *all* as to when she would thus have represented truly. The right thought (on this conception) is one which *would* be true just when she *would* have been representing truly. When is that? Here there is room for Austinian considerations. Ought one who represented as Pia did be held responsible for representing things as any other than they are, or would be if …? When, that is, would it be fair to hold Pia to have fallen anywhere short of the truth? If Sid never inhales, would it be fair/true to

the facts to describe him as a smoker? One understands *what* questions these are only in understanding them as ones about truth's requirements. Abstraction, including Frege's, has its place, but must know it.

REFERENCES

Austin 1950: "Truth", *Proceedings of the Aristotelian Society, Supplementary Volumes*, Vol. 24: 111–128. Reprinted in his *Collected Papers*, 3rd edn., J. O. Urmson and G. J. Warnock (eds.), Oxford: Oxford University Press.

_____ 1962: *How To Do Things With Words*, Cambridge, MA: Harvard University Press.

Frege 1892: "Ueber Sinn und Bedeutung", *Zeitschrift für Philosophie und philosophische Kritik*, NF100, pp. 25–50.

_____ 1897: *Logik*, *Nachgelassene Schriften* (2nd edition), H. Hermes, F. Kambartel, F. Kaulbach eds., Hamburg: Felix Meiner Verlag, 1983, pp. 137–163.

_____ 1918, "Der Gedanke", Beiträge zur deutschen Idealismus 2 1918–1919, pp. 58-77.

P. F.  STRAWSON  &  PAUL  SNOWDON

# BIOGRAPHY

The work of Sir Peter Frederick Strawson, FBA (23 November 1919 – 13 February 2006) spanned philosophy of language, metaphysics, epistemology and the history of philosophy, particularly Kant. He was the Waynflete Professor of Metaphysical Philosophy at the University of Oxford (Magdalen College) from 1968 to 1987. Before that he was appointed as a college lecturer at University College, Oxford in 1947 and became a tutorial fellow the following year until 1968. On his retirement in 1987, he returned to University College and continued working there until shortly before his death. P.F. Strawson was president of the Aristotelian Society from 1969-1970.

"Truth" was originally published in the *Proceedings of the Aristotelian Society, Supplementary Volume XXIV* (1950).

Paul Snowdon has held the post of Grote Professor of Mind and Logic at University College London since 2001 where he is also head of department. His research and publications relate to three main areas: the problem of personal identity (on which he is completing a book manuscript), the philosophy of perception, and the mind-body problem.

# *Truth*

# P. F. STRAWSON

# TRUTH

## P.F. STRAWSON

MR. AUSTIN offers us a purified version of the correspondence theory of truth. On the one hand he disclaims the semanticists' error of supposing that "true" is a predicate of sentences; on the other, the error of supposing that the relation of correspondence is other than purely conventional, the error which models the word on the world or the world on the word. His own theory is, roughly, that to say that a statement is true is to say that a certain speech-episode is related in a certain conventional way to something in the world exclusive of itself. But neither Mr. Austin's account of the two terms of the truth-conferring relation, nor his account of the relation itself, seems to me satisfactory. The correspondence theory requires, not purification, but elimination.

1. *Statements* – It is, of course, indisputable that we use various substantival expressions as grammatical subjects of "true." These are, commonly, noun-phrases like "What he said" or "His statement"; or pronouns or noun-phrases, with a "that"-clause in apposition, *e.g.*, "It ... that *p*" and "The statement that *p*." Austin proposes that we should use "statement" to do general duty for such expressions as these. I have no objection. This will enable us to say, in a philosophically non-committal way, that, in using "true," we are talking about statements. By "saying this in a non-committal way," I mean saying it in a way which does not commit us to any view about the nature of statements so talked about; which does not commit us, for example, to the view that statements so talked about are historic events.

The words "assertion" and "statement" have a parallel and convenient duplicity of sense. "My statement" may be either what I say or my saying it. My saying something is certainly an episode. What I say is not. It is the latter, not the former, we declare to be true. (Speaking the truth is not a manner of speaking: it is saying something true.) When we say "His statement was received with thunderous applause" or "His vehement assertion was followed by a startled silence," we are certainly referring to, characterising, a historic event, and placing it in the context of others. If I say that the same statement was first whispered by John and then bellowed by Peter, uttered first in French and repeated in English, I am plainly still making historical remarks about utterance-

occasions; but the word "statement" has detached itself from reference to any particular speech-episode. The episodes I am talking about are the whisperings, bellowings, utterings and repetitions. The statement is not something that figures in all these episodes. Nor, when I say that the statement is true, as opposed to saying that it was, in these various ways, made, am I talking indirectly about these episodes or any episodes at all. (Saying of a statement that it is true is not related to saying of a speech-episode that it was true as saying of a statement that it was whispered is related to saying of a speech-episode that it was a whisper.) It is futile to ask what thing or event I *am* talking about (over and above the subject-matter of the statement) in declaring a statement to be true; for there is no such thing or event. The word "statement" and the phrase "What he said," like the conjunction "that" followed by a noun clause, are convenient, grammatically substantival, devices, which we employ, on certain occasions, for certain purposes, notably (but not only) the occasions on which we use the word "true." What these occasions are I shall try later to elucidate. To suppose that, whenever we use a singular substantive, we are, or ought to be, using it to refer to something, is an ancient, but no longer a respectable, error.

More plausible than the thesis that in declaring a statement to be true I am talking about a speech-episode is the thesis that in order for me to declare a statement true, there must have occurred, within my knowledge, at least one episode which was a making of that statement. This is largely, but (as Austin sees) not entirely, correct. The occasion of my declaring a statement to be true may be not that someone has made the statement, but that I am envisaging the possibility of someone's making it. For instance, in discussing the merits of the Welfare State, I might say: "It is true that the general health of the community has improved (that $p$), but this is due only to the advance in medical science." It is not necessary that anyone should have said that $p$, in order for this to be a perfectly proper observation. In making it, I am not talking *about* an actual or possible speech-episode. I am myself asserting that $p$, in a certain way, with a certain purpose. I am anticipatorily conceding, in order to neutralize, a possible objection. I forestall someone's making the statement that $p$ by making it myself, with additions. It is of prime importance to distinguish the fact that the use of "true" always glances backwards or forwards to the actual or envisaged making of a statement by someone, from the theory that it is used to characterise such (actual or possible) episodes.

It is not easy to explain the non-episodic and non-committal sense of "statement" in which "statement" = "what is said to be true or false." But, at the risk of being tedious, I shall pursue the subject. For if Austin is right in the suggestion that it is basically of speech-episodes that we predicate "true," it should be possible to "reduce" assertions in which we say of a statement in the non-episodic sense that it is true to assertions in which we are predicating truth of episodes. Austin points out that the same sentence may be used to make different statements. He would no doubt agree that different sentences may be used to make the same statement. I am not thinking only of different languages or synonymous expressions in the same language; but also of such occasions as that on which you say of Jones "He is ill," I say *to* Jones "You are ill" and Jones says "I am ill." Using, not only different sentences, but sentences with different meanings, we all make "the same statement"; and this is the sense of "statement" we need to discuss, since it is, *prima facie*, of statements in this sense that we say that they are true or false (*e.g.*, "What they all said, namely, that Jones was ill, was quite true."). We could say: people make the same statement when the words they use in the situations in which they use them are such that they must (logically) either all be making a true statement or all be making a false statement. But this is to use "true" in the elucidation of "same statement." Or we could say, of the present case: Jones, you and I all make the same statement because, using the words we used in the situation in which we used them, we were all applying the same description to the same person at a certain moment in his history; anyone applying that description to that person (etc.), would be making that statement. Mr. Austin might then wish to analyse (A) "The statement that Jones was ill is true" in some such way as the following: "If anyone has uttered, or were to utter, words such that in the situation in which they are uttered, he is applying to a person the same description as I apply to that person when I now utter the words 'Jones was ill,' then the resulting speech-episode was, or would be, true." It seems plain, however, that nothing but the desire to find a metaphysically irreproachable first term for the correspondence relation could induce anyone to accept this analysis of (A) as an elaborate general hypothetical. It would be a plausible suggestion only if the grammatical subjects of "true" were *commonly* expressions referring to particular, uniquely dateable, speech-episodes. But the simple and obvious fact is that the expressions occurring as such grammatical subjects ("What they

said," "It ... that *p*" and so on) never do, in these contexts, stand for such episodes.[1] *What they said* has no date, though their several sayings of it are dateable. *The statement that p* is not an event, though it had to be made for the first time and made within my knowledge if I am to talk of its truth or falsity. If I endorse Plato's view, wrongly attributing it to Lord Russell ("Russell's view that *p* is quite true"), and am corrected, I have not discovered that I was talking of an event separated by centuries from the one I imagined I was talking of. (Corrected, I may say: "Well it's true, whoever said it.") My *implied* historical judgment is false; that is all.

2. *Facts* – What of the second term of the correspondence relation? For this Mr. Austin uses the following words or phrases : "thing," "event," "situation," "state of affairs," "feature" and "fact." All these are words which should be handled with care. I think that through failing to discriminate sufficiently between them, Mr. Austin (1) encourages the assimilation of facts to things, or (what is approximately the same thing) of stating to referring; (2) misrepresents the use of "true"; and (3) obscures another and more fundamental problem.

In section 3 of his paper, Mr. Austin says, or suggests, that all stating involves both referring ("demonstration") and characterizing ("description"). It is questionable whether all statements do involve both,[2] though it is certain that some do. The following sentences, for example, could all be used to make such statements; *i.e.*, statements in the making of which both the referring and describing functions are performed, the performance of the two functions being approximately (though not exclusively) assignable to different parts of the sentences as uttered:

> The cat has the mange.
> That parrot talks a lot.
> Her escort was a man of medium build, clean-shaven, well-dressed and with a North Country accent.

---

[1] And the cases where such phrases might most plausibly be exhibited as having an episode-referring rôle are precisely those which yield most readily to another treatment; *viz.,* those in which one speaker corroborates, confirms or grants what another has just said (*see* Section 4 below).

[2] *See* Section 5 below. The thesis that all statements involve both demonstration and description is, roughly, the thesis that all statements are, or involve, subject-predicate statements (not excluding relational statements).

In using such sentences to make statements, we refer to a thing or person (object) in order to go on to characterize it: (we demonstrate in order to describe). A *reference* can be correct or incorrect. A *description* can fit, or fail to fit, the thing or person to which it is applied.[3] When we refer correctly, there certainly is a conventionally established relation between the words, so used, and the thing to which we refer. When we describe correctly, there certainly is a conventionally established relation between the words we use in describing and the type of thing or person we describe. These relations, as Mr. Austin emphasizes, are different. An expression used referringly has a different logical rôle from an expression used describingly. They are differently related to the object. And *stating* is different from referring, and different from describing; for it is (in such cases) both these at once. Statement (*some* statement) is reference-cum-description. To avoid cumbersome phrasing, I shall speak henceforward of *parts* of statements (the referring part and the describing part); though parts of statements are no more to be equated with parts of sentences (or parts of speech-episodes) than statements are to be equated with sentences (or speech-episodes).

That (person, thing, etc.) to which the referring part of the statement refers, and which the describing part of the statement fits or fails to fit, is that which the statement is *about*. It is evident that there is nothing else in the world for the statement itself to be related to either in some further way of its own or in either of the different ways in which these different parts of the statement are related to what the statement is about. And it is evident that the demand that there should be such a relatum is logically absurd: a logically fundamental type-mistake. But the demand for something in the world *which makes the statement true* (Mr. Austin's phrase), or *to which the statement corresponds when it is true*, is just this demand. And the answering a theory that to say that a statement is true is to say that a speech-episode is conventionally related in a certain way to such a relatum reproduces the type-error embodied in this demand. For while we certainly say that a statement corresponds to (fits, is borne out by, agrees with) the facts, as a variant on saying that it is true, we *never* say that a statement corresponds to the thing, person, etc., it is about. What "makes the statement" that the cat has mange "true," is not the cat, but the *condition* of the cat, *i.e.,* the fact that the

---

[3] *Cf.* the phrase "He is described as…" What fills the gap is not a sentence (expression which could normally be used to make a statement), but a phrase which could occur as a part of an expression so used.

cat has mange. The only plausible candidate for the position of what (in the world) makes the statement true is the fact it states; but the fact it states is not something in the world.[4] It is not an object; not even (as some have supposed) a complex object consisting of one or more particular elements (constituents, parts) and a universal element (constituent, part). I can (perhaps) hand you, or draw a circle round, or time with a stop-watch the things or incidents that are referred to when a statement is made. Statements are about such objects; but they state facts. Mr. Austin seems to ignore the complete difference of type between, *e.g.,* "fact" and "thing"; to talk as if "fact" were just a very general word (with, unfortunately, some misleading features) for "event," "thing," etc., instead of being (as it is) both wholly different from these, and yet the only possible candidate for the desired non-linguistic correlate of "statement." Roughly: the thing, person, etc., referred to is the material correlate of the referring part of the statement; the quality or property the referent is said to "possess" is the *pseudo*-material correlate of its describing part; and the fact to which the statement "corresponds" is the *pseudo*-material correlate of the statement as a whole.

These points are, of course, reflected in the behaviour of the word "fact" in ordinary language; behaviour which Mr. Austin notes, but by which he is insufficiently warned. "Fact," like "true," "states" and "statement" is wedded to "that"-clauses; and there is nothing unholy about this union. Facts are known, stated, learnt, forgotten, overlooked, commented on, communicated or noticed. (Each of these verbs may be followed by a "that"-clause or a "the fact that"-clause.) Facts are what statements (when true) state; they are not what statements are about. They are not, like things or happenings on the face of the globe, witnessed or heard or seen, broken or overturned, interrupted or prolonged, kicked, destroyed, mended or noisy. Mr. Austin notes the expression "fact that," warns us that it may tempt us to identify facts with true statements and explains its existence by saying that for certain purposes in ordinary life we neglect, or take as irrelevant, the distinction

---

[4] This is not, of course, to deny that there is that in the world which a statement of this kind is about (true or false *of*), which is *referred to* and *described* and which the description fits (if the statement is true) or fails to fit (if it is false). This truism is an inadequate introduction to the task of elucidating, not our use of "true," but a certain general way of using language, a certain type of discourse, *viz.,* the fact-stating type of discourse. What confuses the issue about the use of the word "true" is precisely its entanglement with this much more fundamental and difficult problem. (*See* (ii) of this section.)

between saying something true and the thing or episode of which we are talking. It would indeed be wrong – but not for Mr. Austin's reasons – to identify "fact" and "true statement"; for these expressions have different rôles in our language, as can be seen by the experiment of trying to interchange them in context. Nevertheless their roles – or those of related expressions – overlap. There is no nuance, except of style, between "That's true" and "That's a fact"; nor between "Is it true that...?" and "Is it a fact that...?"[5] But Mr. Austin's reasons for objecting to the identification seem mistaken, as does his explanation of the usage which (he says) tempts us to make it. Because he thinks of a statement as something in the world (a speech-episode) and a fact as something else in the world (what the statement either "corresponds to" or "is about"), he conceives the distinction as of overriding importance in philosophy, though (surprisingly) sometimes negligible for ordinary purposes. But I can conceive of no occasion on which I could possibly be held to be "neglecting or taking as irrelevant" the distinction between, say, my wife's bearing me twins (at midnight) and my saying (ten minutes later) that my wife had borne me twins. On Mr. Austin's thesis, however, my announcing "The fact is that my wife has borne me twins" would be just such an occasion.

Elsewhere in his paper, Mr. Austin expresses the fact that there is no theoretical limit to what could truly be said about things in the world, while there are very definite practical limits to what human beings actually can and do say about them, by the remark that statements "always fit the facts more or less loosely, in different ways for different purposes." But what could fit more perfectly the fact that it is raining than the statement that it is raining? Of course, statements and facts fit. They were made for each other. If you prise the statements off the world you prise the facts off it too; but the world would be none the poorer. (You don't also prise off the world what the statements are about – for this you would need a different kind of lever.)

A symptom of Mr. Austin's uneasiness about facts is his preference for the expressions "situation" and "state of affairs"; expressions of which the character and function are a little less transparent than those

---

[5] I think in general the difference between them is that while the use of "true," as already acknowledged, glances backwards or forwards at an actual or envisaged making of a statement, the use of "fact" does not generally do this though it may do it sometimes. It certainly does not do it in, *e.g.*, the phrase "The fact is that..." which serves rather to prepare us for the unexpected and unwelcome.

of "fact." They are more plausible candidates for inclusion in the world. For while it is true that situations and states of affairs are not seen or heard (any more than facts are), but are rather *summed up* or *taken in at a glance* (phrases which stress the connection with statement and "that"-clause respectively), it is also true that there is a sense of "about" in which we do talk about, do describe, situations and states of affairs. We say, for example, "The international situation is serious" or "This state of affairs lasted from the death of the King till the dissolution of Parliament." In the same sense of "about," we talk about facts; as when we say "I am alarmed by the fact that kitchen expenditure has risen by 50 per cent. in the last year." But whereas "fact" in such usages is linked with a "that"-clause (or connected no less obviously with "statement," as when we "take down the facts" or hand someone the facts on a sheet of paper), "situation" and "state of affairs" stand by themselves, states of affairs are said to have a beginning and an end, and so on. Nevertheless, situations and states of affairs so talked of are (like facts so talked of), abstractions that a logician, if not a grammarian, should be able to see through. Being alarmed by a fact is not like being frightened by a shadow. It is being alarmed because… One of the most economical and pervasive devices of language is the use of substantival expressions to abbreviate, summarize and connect. Having made a series of descriptive statements, I can comprehensively connect with these the remainder of my discourse by the use of such expressions as "this situation" or "this state of affairs"; just as, having produced what I regard as a set of reasons for a certain conclusion I allow myself to draw breath by saying "Since *these things* are so, then…," instead of prefacing the entire story by the conjunction. A situation or state of affairs is, roughly, a set of facts not a set of things.

A point which it is important to notice in view of Mr. Austin's use of these expressions (in sections 3a and 3b of his paper) is that when we *do* "talk about" situations (as opposed to things and persons) the situation we talk about is not, as he seems to think it is, correctly identified with the fact we state (with "what makes the statement true"). If a situation is the "subject" of our statement, then what "makes the statement true" is not the situation, but the fact that the situation has the character it is asserted to have. I think much of the persuasiveness of the phrase "talking about situations" derives from that use of the word on which I have just commented. But if a situation is treated as the "subject" of a statement, then it will not serve as the non-linguistic term, for which Mr. Austin is seeking, of the "relation of correspondence;" and if it is treated

as the non-linguistic term of this relation, it will not serve as the subject of the statement.

Someone might now say "No doubt 'situation,' 'state of affairs,' 'facts' are related in this way to 'that'-clauses and assertive sentences; can serve, in certain ways and for certain purposes, as indefinite stand-ins for specific expressions of these various types. So also is 'thing' related to some nouns; 'event' to some verbs, nouns and sentences; 'quality' to some adjectives; 'relation' to some nouns, verbs and adjectives. Why manifest this prejudice in favour of things and events as alone being parts of the world or its history? Why not situations and facts as well?" The answer to this (implicit in what has gone before) is twofold.

(i) The first part of the answer[6] is that the whole charm of talking of situations, states of affairs or facts as included in, or parts of, the world, consists in thinking of them as things, and groups of things; that the temptation to talk of situations, etc., in the idiom appropriate to talking of things and events is, once this first step is taken, overwhelming. Mr. Austin does not withstand it. He significantly slips in the word "feature" (noses and hills are *features*, of faces and landscapes) as a substitute for "facts." He says that the reason why photographs and maps are not "true" in the way that statements are true is that the relation of a map or a photograph to what it is a map or a photograph of is not wholly (in the first case) and not at all (in the second) a conventional relation. But this is not the only, or the fundamental, reason (The relation between the Prime Minister of England and the phrase "the Prime Minister of England" *is* conventional; but it doesn't make sense to say that someone uttering the phrase out of context is saying something true or false.) The (for present purposes) fundamental reason is that "being a map of" or "being a photograph of" *are* relations, of which the non-photographic, non-cartographical, relata are, say, personal or geographical *entities*. The trouble with correspondence theories of truth is not primarily the tendency to substitute non-conventional relations for what is really a wholly conventional relation. It is the misrepresentation of "correspondence between statement and fact" *as a relation, of any kind, between events or things or groups of things* that is the trouble.

---

[6] Which could be more shortly expressed by saying that if we read "world" (a sadly corrupted word) as "heavens and earth," talk of facts, situations and states of affairs, as "included in" or "parts of" the world is, obviously, metaphorical. The world is the totality of things, not of facts.

Correspondence theorists think of a statement as "describing that which makes it true" (fact, situation, state of affairs) in the way a descriptive predicate may be used to describe, or a referring expression to refer to, a thing.[7]

(ii) The second objection to Mr. Austin's treatment of facts, situations, states of affairs as "parts of the world" which we declare to stand in a certain relation to a statement when we declare that statement true, goes deeper than the preceding one but is, in a sense, its point. Mr. Austin rightly says or implies (section 3) that for some of the purposes for which we use language, there must be conventions correlating the words of our language with what is to be found in the world. Not all the linguistic purposes for which this necessity holds, however, are identical. Orders, as well as information, are conventionally communicated. Suppose "orange" always meant what we mean by "Bring me an orange" and "that orange" always meant what we mean by "Bring me that orange," and, in general, our language contained only sentences in some such way imperative. There would be no less need for a conventional correlation between the word and the world. Nor would there be any less to be found in the world. But those pseudo-entities which *make statements true* would not figure among the non-linguistic correlates. They would no more be found; (they never were found, and never did figure among the non-linguistic correlates). The point is that the word "fact" (and the "set-of-facts" words like "situation" "state of affairs") have, like the words "statement" and "true" themselves, a certain type of word-world-relating discourse (the informative) *built in*

---

[7] Suppose the pieces set on a chessboard, a game in progress. And suppose someone gives, in words, an exhaustive statement of the position of the pieces. Mr. Austin's objection (or one of his objections) to earlier correspondence theories is that they would represent the relation between the description and the board with the pieces on it as like, say, the relation between a newspaper diagram of a chess-problem and a board with the pieces correspondingly arranged. He says, rather, that the relation is a purely conventional one. My objection goes farther. It is that there is no thing or event called "a statement" (though there is the making of the statement) and there is no thing or event called "a fact" or "situation" (though there is the chessboard with the pieces on it) which stand to one another in any, even a purely conventional, relation as the newspaper diagram stands to the board-and-pieces. The facts (situation, state of affairs) cannot, like the chessboard-and-pieces, have coffee spilt on them or be upset by a careless hand. It is because Mr. Austin needs such events and things for his theory that he takes the making of the statement as the statement, and that which the statement is about as the fact which it states.

Events can be dated and things can be located. But the facts which statements (when true) state can be neither dated or located. (Nor can the statements, though the making of them can be.) Are they included in the world?

to them. The occurrence in ordinary discourse of the words "fact" "statement" "true" signalizes the occurrence of this type of discourse; just as the occurrence of the words "order" "obeyed" signalizes the occurrence of another kind of conventional communication (the imperative). If our task were to elucidate the nature of the first type of discourse, it would be futile to attempt to do it in terms of the words "fact," "statement," "true," for these words contain the problem, not its solution. It would, for the same reason, be equally futile to attempt to elucidate any one of these words (in so far as the elucidation of *that* word would be the elucidation of *this* problem) in terms of the others. And it is, indeed, very strange that people have so often proceeded by saying "Well, we're pretty clear what a statement is, aren't we? Now let us settle *the further question*, *viz*., what it is for a statement to be true." This is like "Well, we're clear about what a command is: now what is it for a command to be obeyed?" As if one could divorce statements and commands from the point of making or giving them!

Suppose we had in our language the word "execution" meaning "action which is the carrying out of a command." And suppose someone asked the philosophical question: What is *obedience*? What is it for a command to be *obeyed*? A philosopher might produce the answer: "Obedience is a conventional relation between a command and an execution. A command is obeyed when it corresponds to an execution."

This is the Correspondence Theory of Obedience. It has, perhaps, a little less value as an attempt to elucidate the nature of one type of communication than the Correspondence Theory of Truth has as an attempt to elucidate that of another. In both cases, the words occurring in the solution incorporate the problem. And, of course, this intimate relation between "statement" and "fact" (which is understood when it is seen that they both incorporate this problem) explains why it is that when we seek to explain *truth* on the model of naming or classifying or any other kind of conventional or non-conventional relation between one thing and another, we always find ourselves landed with "fact," "situation," "state of affairs" as the non-linguistic terms of the relation.

But why should the problem of Truth (the problem about our use of "true") be seen as this problem of elucidating the fact-stating type of discourse? The answer is that it shouldn't be; but that the Correspondence Theory can only be fully seen through when it is seen as a barren attempt on this second problem. Of course, a philosopher concerned with the second problem, concerned to elucidate a certain

general type of discourse, must stand back from language and talk about the different ways in which utterances are related to the world (though he must get beyond "correspondence of statement and fact" if his talk is to be fruitful). But – to recur to something I said earlier – the occurrence *in ordinary discourse* of the words "true," "fact," etc., signalizes, without commenting on, the occurrence of a certain way of using language. When we use these words in ordinary life, we are talking within, and not about, a certain frame of discourse; we are precisely not talking about the way in which utterances are, or may be, conventionally related to the world. We are talking about persons and things, but in a way in which we could not talk about them if conditions of certain kinds were not fulfilled. The problem about the use of "true" is to see how this word fits into that frame of discourse. The surest route to the wrong answer is to confuse this problem with the question: What type of discourse is this?[8]

3. *Conventional Correspondence* – It will be clear from the previous paragraph what I think wrong with Mr. Austin's account of the relation itself, as opposed to its terms. In section 4 of his paper he says that, when we declare a statement to be true, the relation between the statement and the world which our declaration "asserts to obtain" is "a purely conventional relation" and "one which we could alter at will." This remark reveals the fundamental confusion of which Mr. Austin is guilty between:

> (*a*) the semantic conditions which must be satisfied for the statement that a certain statement is true to be itself true; and

> (*b*) what is asserted when a certain statement is stated to be true.

Suppose A makes a statement, and B declares A's statement to be true. Then for B's statement to be true, it is, *of course*, necessary that the words used by A in making the statement should stand in a certain conventional (semantical) relationship with the world; and that the "linguistic rules" underlying this relationship should be rules "observed" by both A and B. It should be remarked that these conditions (with the exception of the condition about B's observance of linguistic rules) are

---

[8] A parallel mistake would be to think that in our ordinary use (as opposed to a philosopher's use) of the word "quality," we were talking about people's uses of words; on the ground (correct in itself) that this word would have no use but for the occurrence of a certain general way of using words.

equally necessary conditions of A's having made a true statement in using the words he used. *It is no more and no less absurd to suggest that B, in making his statement, asserts that these semantic conditions are fulfilled than it is to suggest that A, in making his statement, asserts that these semantic conditions are fulfilled* (*i.e.*, that we can never use words without mentioning them). If Mr. Austin is right in suggesting that to say that a statement is true is to say that "the historic state of affairs to which it [i.e., for Mr. Austin, the episode of making it] is correlated by the demonstrative conventions (the one it 'refers to') is of a type with which the sentence used in making the statement is correlated by the descriptive conventions," then (and this is shown quite clearly by his saying that the relation we assert to obtain is a "purely conventional one" which "could be altered at will") in declaring a statement to be true, we are either:

> (*a*) talking about the meanings of the words used by the speaker whose making of the statement is the occasion for our use of "true" (*i.e.*, profiting by the occasion to give semantic rules); or

> (*b*) saying that the speaker has used correctly the words he did use.

It is *patently* false that we are doing either of these things. Certainly, we use the word "true" when the semantic conditions described by Austin[9] are fulfilled; but we do not, in using the word, *state* that they are fulfilled. (And this, incidentally, is the answer to the question with which Mr. Austin concludes his paper.) The damage is done (the two problems distinguished at the end of the previous section confused) by asking the

---

[9] In what, owing to his use of the words "statement" "fact" "situation," etc., is a misleading form. The quoted account of the conditions of truthful datement is more nearly appropriate as an account of the conditions of correct descriptive reference. Suppose, in a room with a bird in a cage, I say "That parrot is very talkative." Then my use of the referring expression ("That parrot") with which my sentence begins is correct when the token-object (bird) with which my token-expression (event) is correlated by the conventions of demonstration is of a kind with which the type-expression is correlated by the conventions of description. Here we do have an event and a thing and a (type-mediated) conventional relation between them. If someone corrects me, saying "That's not a parrot; it's a cockatoo," he may be correcting either a linguistic or a factual error on my part. (The question of which he is doing is the question of whether I would have stuck to my story on a closer examination of the bird.) Only in the former case is he declaring a certain semantic condition to be unfulfilled. In the latter case, he is talking about the bird. He asserts that it is a cockatoo and not a parrot. This he could have done whether I had spoken or not. He also *corrects* me, which he could not have done if I had not spoken.

question: *When* do we use the word "true"? instead of the question: *How* do we use the word "true"?

Someone says: "It's true that French Governments rarely last more than a few months, but the electoral system is responsible for that." Is the fact he states in the first part of his sentence alterable by changing the conventions of language? It is not.

4. *Uses of "that"-clauses; and of "statement," "true," "fact," "exaggerated," etc.* – (*a*) There are many ways of making an assertion about a thing, X, besides the bare use of the sentence-pattern "X is Y." Many of these involve the use of "that"-clauses. For example:–

| | |
|---|---|
| How often shall I have to tell you | |
| Today I learnt | |
| It is surprising | |
| The fact is | |
| I have just been reminded of the fact | that X is Y. |
| It is indisputable | |
| It is true | |
| It is established beyond question | |

These are all ways of asserting, in very different context and circumstances, that X is Y. [10] Some of them involve autobiographical assertions as well; others do not. In the grammatical sense already conceded, all of them are "about" facts or statements. In no other sense is any of them about either, though some of them carry *implications* about the *making* of statements.

(*b*) There are many different circumstances in which the simple sentence-pattern "X is Y" may be used to do things which are not merely stating (though they all involve stating) that X is Y. In uttering words of this simple pattern we may be encouraging, reproving or warning someone; reminding someone; answering, or replying to, someone; denying what someone has said; confirming, granting, corroborating, agreeing with, admitting what someone has said. Which

---

[10] One might prefer to say that in some of these cases one was asserting only *by implication* that X is Y; though it seems to me more probable that in all these cases we should say, of the speaker, not "What he said implied that X is Y," but " He *said* that X was Y."

of these, if any, we are doing depends on the circumstances in which, using this simple sentence-pattern, we assert that X is Y.

(*c*) In many of the cases in which we are doing something besides merely stating that X is Y, we have available, for use in suitable contexts, certain abbreviatory devices which enable us to state that X is Y (to make our denial, answer, admission or whatnot) *without* using the sentence-pattern "X is Y." Thus, if someone asks us "Is X Y?", we may state (in the way of reply) that X is Y by saying "Yes." If someone says "X is Y," we may state (in the way of denial) that X is not Y, by saying "It is not" or by saying "That's not true"; or we may state (in the way of corroboration, agreement, granting, etc.) that X is Y by saying "It is indeed" or "That is true." In all these cases (of reply, denial and agreement) the context of our utterance, as well as the words we use, must be taken into account if it is to be clear what we are asserting, viz., that X is (or is not) Y. It seems to me plain that in these cases "true" and "not true" (we rarely use "false") are functioning as abbreviatory statement – devices of the same general kind as the others quoted. And it seems also plain that the *only* difference between these devices which might tempt us, while saying of some ("Yes," "It is indeed," "It is not") that, in using them, we were talking about X, to say of others ("That's true," "That's not true") that, in using them, we were talking about something quite different, *viz.*, the utterance which was the occasion for our use of these devices, is their difference in grammatical structure, *i.e.,* the fact that "true" occurs as a grammatical predicate.[11] (It is obviously not a predicate of X.) If Mr. Austin's thesis, that in using the word "true" we make an assertion about a statement, were no more than the thesis that the word "true" occurs as a grammatical predicate, with, as grammatical subjects, such words and phrases as "That," "What he said," "His statement," etc., then, of course, it would be indisputable. It is plain, however, that he means more than this, and I have already produced my objections to the more that he means.

(d) It will be clear that, in common with Mr. Austin, I reject the thesis that the phrase "is true" is logically superfluous, together with the thesis that to say that a proposition is true is *just* to assert it and to say

---

[11] Compare also the English habit of making a statement followed by an interrogative appeal for agreement in such forms as "isn't it?", "doesn't he? " etc., with the corresponding German and Italian idioms, "Nicht wahr?", "non è vero?" There is surely no significant difference between the phrases which do not employ the word for "true" and those which do: they all appeal for agreement in the same way.

that it is false is *just* to assert its contradictory. "True" and "not true" have jobs of their own to do, *some*, but by no means all, of which I have characterized above. In using them, we are not *just* asserting that X is Y or that X is not Y. We are asserting this in a way in which we could not assert it unless certain conditions were fulfilled; we may also be granting, denying, confirming, etc. It will be clear also that the rejection of these two theses does not entail acceptance of Mr. Austin's thesis that in using "true" we are making an assertion about a statement. Nor does it entail the rejection of the thesis which Mr. Austin (in Section 4 of his paper) couples with these two, *viz*., the thesis that to say that an assertion is true is not to make any further *assertion* at all. This thesis holds for many uses, but requires modification for others.

(e) The occasions for using "true" mentioned so far in this section are evidently not the only occasions of its use. There is, for example, the generally concessive employment of "It is true that *p*…", which it is difficult to see how Mr. Austin could accommodate. All these occasions have, however, a certain contextual immediacy which is obviously absent when we utter such sentences as "What John said yesterday is quite true" and "What La Rochefoucauld said about friendship is true." Here the context of our utterance does not identify for us the statement we are talking about (in the philosophically non-committal sense in which *we are* "talking about statements" when we use the word "true"), and so we use a descriptive phrase to do the job. But the descriptive phrase does not identify an event; though the statement we make carries the implication (in some sense of "implication") that there occurred an event which was John's making yesterday (or Rochefoucauld's making sometime) the statement that *p* (*i.e.*, the statement we declare to be true). We are certainly not telling our audience that the event occurred, *e.g.*, that John made the statement that *p*, for (i) we do not state, either by way of quotation or otherwise, what it was that John said yesterday, and (ii) our utterance achieves its main purpose (that of making, by way of confirmation or endorsement, the statement that *p*) only if our audience already knows that John yesterday made the statement that *p*. The abbreviatory function of "true" in cases such as these becomes clearer if we compare them with what we say in the case where (i) we want to assert that *p*; (ii) we want to indicate (or display our knowledge that) an event occurred which was John's making yesterday the statement that *p*; (iii) we believe our audience ignorant or forgetful of the fact that John said yesterday that *p*. We then use the formula "As John said yesterday, *p*" or "It is true, as John said yesterday, that *p*," or "What John said yesterday, namely that *p*, is true." (Of course the words represented by

the letter *p*, which we use, may be – sometimes, if we are to make the same statement, must be – different from the words which John used.) Sometimes, to embarrass, or test, our audience, we use, in cases where the third of these conditions is fulfilled, the formula appropriate to its non-fulfilment, *viz*., "What John said yesterday is true."

(*f*) In criticism of my view of truth put forward in *Analysis*,[12] and presumably in support of his own thesis that "true" is used to assert that a certain relation obtains between a speech-episode and something in the world exclusive of that episode, Mr. Austin makes, in Section 7 of his paper, the following point. He says: "Mr. Strawson seems to confine himself to the case when I say "Your statement is true" or something similar – but what of the case when you state that S and I say nothing, but *look and see* that your statement is true?" The point of the objection is, I suppose, that since I *say* nothing, I cannot be making any performatory use of "true"; yet I can see *that* your statement is true. The example, however, seems to have a force precisely contrary to what Mr. Austin intended. Of course, "true" has a different rôle in "X sees that Y's statement is true" from its rôle in "Y's statement is true." What is this rôle? Austin says in my hearing "There is a cat on the mat" and I look and see a cat on the mat. Someone (Z) reports: "Strawson saw that Austin's statement was true." What is he reporting? He is reporting that I have seen a cat on the mat; but he is reporting this in a way in which he could not report it except in certain circumstances, *viz*., in the circumstances of Austin's having said in my hearing that there was a cat on the mat. Z's remark also carries the implication that Austin made a statement, but cannot be regarded as *reporting* this by implication since it fulfils its main purpose only if the audience already knows that Austin made a statement and what statement he made; and the implication (which *can* be regarded as an implied report) that I heard and understood what Austin said.[13] The man who looks and sees that the statement that there is a cat on the mat is true, sees no more and no less than the man who looks and sees that there is a cat on the mat, or the man who looks and sees that there is *indeed* a cat on the mat. But the *settings* of the first and third cases may be different from that of the second.

---

[12] Vol. 9, No. 6, June, 1949.
[13] If *I* report: "I see that Austin's statement is true," this is simply a first-hand corroborative report that there is a cat on the mat, made in a way in which it could not be made except in these circumstances.

This example has value, however. It emphasizes the importance of the concept of the "occasion" on which we may make use of the assertive device which is the subject of this symposium (the word "true"); and minimizes (what I was inclined to over-emphasize) the performatory character of our uses of it.

(*g*) Mr. Austin stresses the differences between negation and falsity; rightly, in so far as to do so is to stress the difference (of occasion and context) between asserting that X is not Y and denying the assertion that X is Y. He also exaggerates the difference; for, if I have taken the point of his example, he suggests that there are cases in which "X is not Y" is inappropriate to a situation in which, if anyone stated that X was Y, it would be correct to say that the statement that X was Y was false. These are cases where the question of whether X is or is not Y does not arise (where the conditions of its arising are not fulfilled). They are equally, it seems to me, cases when the question of the truth or falsity of the statement that X is Y does not arise.

(*h*) A qualification of my general thesis, that in using "true" and "untrue" we are not talking about a speech-episode, is required to allow for those cases where our interest is not primarily in what the speaker asserts, but in the speaker's asserting it, in, say, the fact of his having *told the truth* rather than in the fact which he reported in doing so. (We may, of course, be interested in both; or our interest in a man's evident truthfulness on one occasion may be due to our concern with the degree of his reliability on others.)

But this case calls for no special analysis and presents no handle to any theorist of truth; for to use "true" in this way is simply to characterize a certain *event* as *the making*, by someone, of a true statement. The problem of analysis remains.

(*i*) Mr. Austin says that we shall find it easier to be clear about "true" if we consider other adjectives "in the same class," such as "exaggerated," "vague," "rough," "misleading," "general," "too concise." I do not think these words *are* in quite the same class as "true" and "false." In any language in which statements can be made at all, it must be possible to make true and false statements. But statements can suffer from the further defects Mr. Austin mentions only when language has attained a certain richness. Imagine one of Mr. Austin's rudimentary languages with "single words" for "complex situations" of totally different kinds. One could make true or false statements; but not

statements which were exaggerated, over-concise, too general or rather rough. And even given a language as rich as you please, whereas all statements made in it could be true or false, not all statements could be exaggerated. When can we say that the statement that $p$ is exaggerated? One of the conditions is this: that, if the sentence $S_1$ is used to make the statement that p, there should be some sentence $S_2$ (which could be used to make the statement that $q$) such that $S_1$ and $S_2$ are related somewhat as "There were 200 people there" is related to "There were 100 people there." (To the remark "We got married yesterday," you cannot, except as a joke, reply: "You're exaggerating.")

Mr. Austin's belief, then, that the word "exaggerated" stands for a relation between a statement and something in the world exclusive of the statement, would at least be an over-simplification, even if it were not objectionable in other ways. But it is objectionable in other ways. The difficulties about statement and fact recur; and the difficulties about the relation. Mr. Austin would not want to say that the relation between an exaggerated statement and the world was like that between a glove and a hand too small for it. He would say that the relation was a conventional one. But the fact that the statement that $p$ is exaggerated is not in any sense a conventional fact. (It is, perhaps, the fact that there were 1,200 people there and not 2,000.) If a man says: "There were at least 2,000 people there," you may reply (A) "No, there were not so many (far more)," or you may reply (B) "That's an exaggeration (understatement)." (A) and (B) say the same thing. Look at the situation more closely. In saying (A), you are not merely asserting that there were fewer than 2,000 people there: you are also correcting the first speaker, and correcting him in a certain general way, which you could not have done if he had not spoken as he did, though you could merely have asserted that there were fewer than 2,000 there without his having spoken. Notice also that what is being asserted by the use of (A) – that there were fewer than 2,000 there – cannot be understood without taking into account the original remark which was the occasion for (A). (A) has both contextually-assertive and performatory features. (B) has the same features, and does the same job as (A), but more concisely and with greater contextual reliance.

Not all the words taken by Austin as likely to help us to be clear about "true" are in the same class as one another. "Exaggerated" is, of those he mentions, the one most relevant to his thesis; but has been seen to yield to my treatment. Being "over-concise" and "too general" are not ways of being "not quite true." These obviously relate to the specific

purposes of specific makings of statements; to the unsatisfied wishes of specific audiences. No alteration in things in the world, nor any magical replaying of the course of events, could bring statements so condemned into line, in the way that an "exaggerated assessment" of the height of a building could be brought into line by inorganic growth. Whether the statement (that *p*) is true or false is a matter of the way things are (of whether *p*); whether a statement is exaggerated (if the question arises – which depends on the type of statement and the possibilities of the language) is a matter of the way things are (*e.g.,* of whether or not there were fewer than 2,000 there). But whether a statement is over-concise[14] or too general depends on what the hearer wants to know. The world does not demand to be described with one degree of detail rather than another.

5. *The scope of "statement," "true," "false" and "fact"* – Commands and questions, obviously do not claim to be statements of fact: they are not true or false. In Section 6 of his paper, Mr. Austin reminds us that there are many expressions neither interrogative nor imperative in form which we use for other purposes than that of reportage or forecast. From our employment of these expressions he recommends that we withhold (suspects that we do, in practice, largely withhold) the appellation "stating facts," the words "true" and "false." Philosophers, even in the sphere of language are not legislators; but I have no wish to challenge the restriction, in some philosophical contexts, of the words "statement," "true," "false," to what I have myself earlier called the "fact-stating" type of discourse.

What troubles me more is Mr. Austin's own incipient analysis of this type of discourse. It seems to me such as to force him to carry the restriction further than he wishes or intends. And here there are two points which, though connected, need to be distinguished. First, there are difficulties besetting the relational theory of truth as such; second, there is the persistence of these difficulties in a different form when this "theory of truth" is revealed as, rather, an incipient analysis of the statement-making use of language.

---

[14] "Concise" is perhaps less often used of what a man says than of the way he says it (*e.g.*, "concisely put," "concisely expressed," "a concise formulation"). A may take 500 words to say what B says in 200. Then I shall say that B's formulation was more concise than A's, meaning simply that he used fewer words.

First then, facts of the cat-on-the-mat-type are the favoured species for adherents of Mr. Austin's type of view. For here we have one thing (one chunk of reality) sitting on another: we can (if we are prepared to commit the errors commented on in Section (2) above) regard the two together as forming a single chunk, if we like, and call it a fact or state of affairs. The view may then seem relatively plausible that to say that the statement (made by me to you) that the cat is on the mat is true is to say that the three-dimensional state of affairs with which the episode of my making the statement is correlated by the demonstrative conventions is of a type with which the sentence I use is correlated by the descriptive conventions. Other species of fact, however, have long been known to present more difficulty: the fact that the cat is not on the mat, for example, or the fact that there are white cats, or that cats persecute mice, or that if you give my cat an egg, it will smash it and eat the contents. Consider the simplest of these cases, that involving negation. With what type of state-of-affairs (chunk of reality) is the sentence "The cat is not on the mat" correlated by conventions of description? With a mat *simpliciter*? With a dog on a mat? With a cat up a tree? The amendment of Mr. Austin's view to which one might be tempted for negative statements (*i.e.,* "S is true" = "The state of affairs to which S is correlated by the demonstrative conventions is not of a type with which the *affirmative form of* S is correlated by the descriptive conventions") destroys the simplicity of the story by creating the need for a different sense of "true" when we discuss negative statements. And worse is to follow. Not all statements employ conventions of demonstration. Existential statements don't, nor do statements of (even relatively) unrestricted generality. Are we to deny that these are statements, or create a further sense of "true"? And what has become of the non-linguistic correlate, the chunk of reality? Is this, in the case of existential or general statements, the entire world? Or, in the case of negatively existential statements, an ubiquitous non-presence?

As objections to a correspondence theory of truth, these are familiar points; though to advance them as such is to concede too much to the theory. What makes them of interest is their power to reveal how such a theory, in addition to its intrinsic defects, embodies too narrow a conception of the fact-stating use of language. Mr. Austin's description of the conditions under which a statement is true, regarded as an analysis of the fact-stating use, applies only to affirmative subject-predicate statements, *i.e.,* to statements in making which we refer to some one or more localized thing or group of things, event or set of events, and characterize it or them in some positive way (identify the

object or objects and affix the label). It does not apply to negative, general and existential statements nor, straightforwardly, to hypothetical and disjunctive statements. I agree that any language capable of the fact-stating use must have some devices for performing the function to which Mr. Austin exclusively directs his attention, and that other types of statements of fact can be understood only in relation to this type. But the other types *are* other types. For example, the word "not" can usefully be regarded as a kind of crystallizing-out of something *implicit* in all use of descriptive language (since no predicate would have any descriptive force if it were compatible with everything). But from this it does not follow that negation (*i.e.,* the *explicit* exclusion of some characteristic) is a kind of affirmation, that negative statements are properly discussed in the language appropriate to affirmative statements. Or take the case of existential statements. Here one needs to distinguish two kinds of demonstration or reference. There is, first, the kind whereby we enable our hearer to identify the thing or person or event or set of these which we then go on to characterize in some way. There is, second, the kind by which we simply indicate a locality. The first ("*Tabby* has the mange") answers the question "Who, which one, what) are you talking about?" The second ("*There's* a cat") the question "Where?" It is plain that no part of an existential statement performs the first function; though Austin's account of reference-cum-description is appropriate to reference of this kind rather than to that of the other. It is clear also that a good many existential statements do not answer the question "Where?" though they may license the enquiry. The difference between various types of statement, and their mutual relations, is a matter for careful description. Nothing is gained by lumping them all together under a description appropriate only to one, even though it be the basic, type.

6. *Conclusion* – My central objection to Mr. Austin's thesis is this. He describes the conditions which must obtain if we are correctly to declare a statement true. His detailed description of these conditions is, with reservations, correct as far as it goes, though in several respects too narrow. The central mistake is to suppose that in using the word "true" we are asserting such conditions to obtain. That this is a mistake is shown by the detailed examination of the behaviour of such words as "statement," "fact," etc., and of "true" itself, and by the examination of various different types of statement. This also reveals some of the ways in which "true" actually functions as an assertive device. What supremely confuses the issue is the failure to distinguish between the task of elucidating the nature of a certain type of communication (the

empirically informative) from the problem of the actual functioning of the word "true" within the framework of that type of communication.

*Strawson's Truth*

PAUL SNOWDON

# STRAWSON'S TRUTH

## PAUL SNOWDON

STRAWSON'S 'Truth' is his response to Austin's paper of the same title, both of which were presented to the Aristotelian Society in the summer of 1950. Because it is primarily a critical response Strawson's paper is not quite typical of his style. Usually a paper by Strawson is a continuous and developing argument for a philosophical theory. But in this case he comments on Austin's approach in a series of remarks, which to some extent are relatively separate. However, Strawson has two other more general goals besides considering or criticising Austin. The first is to criticise a more general theory which Strawson calls the Correspondence Theory of Truth, Austin's presumably being only one version of it. The second is to articulate to some extent and display the virtues of what is often called the Redundancy Theory of Truth, originally proposed by Ramsey, which fundamentally Strawson was convinced by. Strawson was, though, unhappy about that familiar name.

Now, because Strawson's paper is, as one might say, basically 'reactive', it cannot be properly considered in isolation from Austin's paper. What then is Austin doing in his paper? I think that we can see it also as having three main aims. The first is to identify the things which are true (Austin calls them 'statements') and to specify the conditions under which they are true. This amounts to his theory of truth. Austin of course wittily insisted that the focus of analysis should be the adjective 'true' rather than the noun 'truth', but we are hardly being unfair to him by calling what he offers a theory of *truth*. Second, towards the end of his paper Austin criticises something like the Ramsey approach, and some additions that Strawson himself had proposed to it in an earlier paper. But, third, as I read it, Austin is trying to convey a message about truth and what we might call the 'philosophy of truth'. It is not an easy message to summarise but I think that one can say that the message, or part of the message, is, or includes, the proposal that analysing truth is not especially difficult, nor is truth as important or interesting as philosophers usually suppose. According to Austin the real questions include such things as; how do the 'symbols' used in map-making differ from those used in statement-making? He compares focussing on truth to focussing on freedom, which is according to him far too general a notion to analyse profitably. He proposes, too, that truth is not really the basic goal of assertion. Austin in fact was attempting in a way to lead philosophers *away* from the analysis of truth. In this Austin is engaged in conveying a negative attitude to standard philosophy, an activity that for him was almost inseparable from doing philosophy at all. Interestingly, I think it is true to say that Strawson does not himself really engage head on with this aspect of Austin's paper, although he picks up some issues connected to it.

Now, whatever one's reactions to these central proposals of Austin's fairly short paper, it is hard, I feel, to escape the feeling that the paper does not represent Austin at his best. Quite a number of times he makes points that look questionable or arbitrary. For example, in his discussion of whether beliefs are the basic bearers of truth he simply remarks that 'if, as some also say, a belief is 'of the nature of a picture' then it is of the nature of what cannot be true.' [1] To which one feels like responding; whoever thinks beliefs are pictures? Where does that idea so much as come from? Again, when considering the proposal that propositions are the basic truth bearers Austin remarks that 'a proposition in law or in geometry is something portentous' and seems to imply that that is a defining characteristic of a proposition.[2] Surely Austin was familiar with employment of that term outside those contexts, or perhaps with its employment by philosophers. Although I cannot substantiate the claim here I believe that Austin's paper offers a number of examples of this. This perhaps links with the verdict that the encounter between Strawson and Austin was 'won' by Strawson. Such infelicities would not have escaped the assembled philosophers listening to the encounter. Of course, that Austin committed various mistakes should not be taken to mean that his main claims are wrong.

Austin's theory of truth is as follows; the primary bearers of truth are statement, which are, or seem to be, sentences as used by a certain person on a certain occasion, and a statement is true when the historic state of affairs to which it is correlated by the demonstrative conventions is of a type with which the sentence used in making it is correlated by the descriptive conventions'.[3] This whole structure acted on Strawson rather as a red rag to a bull. He objects (in section I) to Austin's notion of a statement (and its role), and then (in section II) to the other end of the relation, facts or states of affairs, or situations, and then (in section III) to Austin's account of the so called conventions which define the relation and their role in his analysis of 'true'.

In his discussion of Austin's identification of the truth bearers as statements, Strawson suggests, very plausibly it seems to me, that Austin's understanding of 'statement', as, it seems, a token use of some words, while perfectly legitimate as one reading of 'statement', does not really pick out what, in many contexts when we are ascribing truth, we can be taken to be talking about. Thus, if I say that his statement is true, what I say could be re-expressed by saying that what he said is true, and what he said is clearly not the words he used in making the statement. In this sense of what he said, it is obvious that what Jo said and what Margaret said can be, as we would say, the same thing, but their statings of it are not the same thing.

---

[1] Austin (1950) p. 86. Page references are to Austin (1961).
[2] Austin (1950) p. 86.
[3] For statements see Austin (1950) pp 87 – 89, for the conventions see p.90.

Conceding the validity of Strawson's point here leaves at least two interesting questions. The first is about Austin. Can Austin's analysis be maintained or defended if he were to drop his thesis that the basic bearers of truth are statements *in his sense*? The problem is that central to Austin's approach is the reference to two types of conventions which relate to language, or language use, and they need, somewhere in the analysis, a linguistic hook to attach to. But the non-linguistic use of statement does not immediately provide that.

The second question relates to Strawson. He says that 'it is not easy to explain the non-episodic and non-committal sense of statement', by which he means its use as expressing what various people might have said, and which according to him is what truth ascriptions deal with.[4] The question is whether Strawson can really think there are truth bearers, if ultimately talk of truth is simply a linguistic device for re-assertion. Really, on that conception, since nothing is ascribed, but rather a claim is reaffirmed, there is nothing for anything to bear, hence no need for truth bearers.

This question arises from a suspicion that there might be a contradiction in holding both that there is no such property as truth (which the 'redundancy theory' is committed to) and that statements are the truth-bearers. Strawson is, though, alive to this issue. When talking about the expression 'statement' he describes it as a 'convenient, grammatically substantive, device' and immediately remarks that there is no necessity to suppose that we are using it to 'refer to something'. His view is, then, that the employment of the noun 'statement' need not commit us to *things* which bear truth.[5]

We can say, then, that Strawson's attitude to statement-talk and his acceptance of a Ramsey-style view of truth fit together in 1950. It is not clear that they continued to do so. By 1998 Strawson is talking of a statements as a type of thing, namely 'an intensional abstract item' and he is talking of truth as a genuine property. It therefore become unclear what his attitude is, by then, to Ramsey's proposal.[6]

The points, far too rich to summarise, that Strawson makes in section II, a section which is 9 pages long, certainly, it seems to me, raise serious questions about Austin's conception of what we might call the world end of the 'correspondence' relation. Austin talks of facts, states of affairs, and situations, but when he states his account the preferred expression is 'state of affairs'. Austin's own remarks about facts seem designed to discourage employing that notion in the theory of truth. Using the notion of a state of affairs Austin formulates his proposal thus; 'A statement is said to be true when the historic state of affairs to which it

---

[4] Strawson (1950) p.192. Page references are to the reprint of the paper in Strawson (1971).
[5] The quotations in this paragraph come from p. 191 of Strawson (1950).
[6] This is how Strawson speaks in Strawson (1998).

is correlated by the demonstrative conventions (the one to which it refers)is of  type with which the sentence used in making it is correlated by the descriptive conventions.'[7] There are two features of this to which I wish to draw attention. 1] Austin clearly wants to say that states of affairs are 'found in the world', that is they are worldly things. However, despite wanting to say this and making them central to his account Austin devotes hardly any attention at all as to what they really are. Reading Austin with Strawson's very careful discussion of the expression 'state of affairs' in mind brings home how insouciant Austin himself is. 2] Taking, as he does, states of affairs as the basic elements Austin has to employ that notion in his description of the conventions of language, since fundamentally the theory captures what truth is in terms of linguistic conventions and the world. But amazingly Austin makes hardly any effort to show that it is correct to analyse the basic conventions of language in terms of states of affairs.  Once that issue is seriously posed it is hard not to feel that it is not a very natural proposal.

Strawson proposes something like a redundancy theory of 'facts'. That is, 'It is a fact that p' is simply equivalent to 'p'. That would imply that facts are not things in the world; talk of facts is simply a way of talking about the proper world occupying things which are really there. Strawson also argues that talk of states of affairs and situations is equally no more about special things *in the world* than is talk about facts. I think it can be said that Austin did not provide a properly investigated employment of these expressions. What needs to be asked though about Strawson's treatment is whether he properly shows his redundancy type theory of facts is correct. An aspect of this is that even if facts are not properly to be thought of as within the spatio-temporal world, there, as it were, to be tripped over, maybe they are entities of another, abstract, sort, and if so, maybe they can also figure somewhere in a theory of truth. The question here is how far Strawson shuts out the Correspondence theory. I think that Strawson's response at this point in the debate would be that correspondence theorists need the world end of the correspondence relation be something properly *in the world*, and moving facts or states of affairs into the abstract realm prohibits them fulfilling that role. In effect, I believe, that is Strawson's argument against the general Correspondence Theory.

Having so far made a very strong case for saying that Austin had selected the wrong interpretation of 'statement' and had not properly thought through the significance of his talk of states of affairs, etc, Strawson turns, in section III, to Austin's conception of the two kinds of conventions that an account of truth needs to attend to. Amazingly, Strawson is quite gentle with Austin's conception of two types of conventions. I suspect that this reflects Strawson's acceptance of the deep importance of the contrast between reference and description (or characterisation) in thinking about the  understanding of language. He

---

[7] Austin (1950) p. 90.

wants to allow *something like* Austin's two elements in an account of language. Fifteen years later, in his paper 'Truth; a reconsideration of Austin's views', Strawson brings out that Austin's talk of two types of conventions is very hard to make sense of. For example, presumably, the sentence 'That cat is sleeping' is descriptively correlated with the type of situation of an indicated cat sleeping; but then when I use it in on a particular occasion what other convention is involved? Rather, I simply indicate a particular cat. In 'Truth', though, Strawson chooses not to develop radical criticisms of Austin's talk of conventions. What he says, though, is that it is simply obvious, or 'patent', that when we ascribe truth to something we are *not* saying anything about linguistic conventions. As Strawson puts it, it may be that Austin has said something which is close to being correct and perhaps informative about *when* a remark is true, but that is not what we are asserting in saying that it is true.

I want to make four remarks about this central thought in Strawson's paper. The first is that Strawson's style of argument is a style he used on other occasions. Basically, Strawson trusted his sense as an ordinary and reflective speaker as to what he was saying when employing certain constructions, be it using the word 'true', or using definite descriptions, as a basis for rejecting a philosophical analysis of that talk that was under consideration. He does not solely look for implications of the theory that might worry one; he held that it is simply obvious to ordinary speakers that the proposal, be it Austin's or Russell's, was wrong. The second remark is that there is something surprising in Strawson's concession that Austin's proposal may come close to being correct about *when* a statement is true. It is natural to feel that Strawson's rather savage handling of some of the elements in the analysis would have led him to deny Austin had even got the 'when' question right. In Strawson (1965) that is precisely what he very effectively questions. The third remark amounts to a question; is Strawson really entitled to be sure that in the relevant philosophical sense of what we are saying, that we are not saying what Austin proposes? I am myself not offering an answer to that question, but I want to note two things. First, Strawson himself accepted analyses of ordinary concepts or terms– for example, the concept of perception – of which it cannot be said, I am inclined to think, that the ordinary speaker would feel it represents what he is saying when employing the relevant term. Second, the goal of philosophical analysis is often specified in terms of a priori determinable necessary and sufficient conditions, a conception that the ordinary speaker, even a highly intelligent and subtle one, can hardly be authoritative about. The fourth point, I hope, is in Strawson's favour. It seems to me that the type of position that Strawson was developing is one that cannot be ruled out on general grounds. It is one thing to say that we employ a certain part of speech *when* certain conditions are fulfilled and quite another to claim that the role in the language of that part of speech is to say, even in an extended sense, that those conditions are fulfilled.

There is much more in Strawson that deserves attention, but I want to conclude by engaging with one other aspect of Strawson's paper. Although Strawson does not say this he basically agreed with Ramsey's so called Redundancy theory. Now, Strawson did not try to remove one problem for that view, which is how the analysis applies to use of 'true' where there is no attached claim that is being affirmed, as in 'What he said to you is true'. But he felt in 1950 that the basic idea could be supplemented by talking of speech acts that we perform when using 'true' which we would not perform if we just affirmed the attached claim. Thus, I can count as conceding that p if I say 'It is true that p', whereas simply affirming that p is not to concede anything. Strawson's own major criticism of his paper is that this attempt to add to Ramsey by 'taking a leaf out of Austin's own book' is not helpful. It obscures the central point of Ramsey's theory and it brings in facts about speech acts that are too unsystematic to clarify the use of 'true'. With Strawson's major retrospective criticism we should, surely, agree. Strawson did not ever, I think, abandon his commitment to Ramsey, and it would not be insane to conjecture that his own subsequent criticisms of the Davidsonian programme in part reflected what might be called an implication of Ramsey, which is that if 'true' is basically a transparent device for reassertion it cannot really be introducing something substantive enough to be the central notion in the theory of meaning.

Strawson's thorough, subtle and professional paper surely throws a number of spanners into Austin's works. It remains unclear though quite what power the spanner that Strawson thought was the biggest has.[8]

---

REFERENCES

Austin J. L. (1950) 'Truth', reprinted in Austin (1961) pp.85 – 101.

Austin J. L ( 1961) *Philosophical Papers* (Oxford, Oxford University Press).

(ed.) Hahn L. E. (1998) *The Philosophy of P. F. Strawson* (Illinois, Open Court).

Strawson P. F. (1950) 'Truth', reprinted in Strawson (1971) pp. 190 – 213.

Strawson P. F. (1965) 'Truth; a Reconsideration of Austin's View', reprinted in Strawson (1971) pp.234 – 249.

Strawson P. F. (1971) Logico-Linguistic Papers (London, Methuen & Co Ltd.).

Strawson P. F. (1998) 'Reply to Searle' in (ed) Hahn L. E. (1998).

MICHAEL DUMMETT & IAN RUMFITT

# BIOGRAPHY

Sir Michael Anthony Eardley Dummett, FBA, D.Litt (27 June 1925 – 27 December 2011) was, until 1992, Wykeham Professor of Logic at the University of Oxford. He wrote on the history of analytic philosophy, most notably as an interpreter of Frege, and has made original contributions to the subject, particularly in the philosophies of mathematics, logic, language and metaphysics. He was best known for his work on truth and meaning and their implications for the debates between realism and anti-realism, a term he helped popularize.

"Truth" was originally published in the *Proceedings of the Aristotelian Society, Supplementary Volume LIX* (1959).

Ian Rumfitt studied philosophy at Balliol College, Oxford, and at Princeton University, and has taught it at Keele University, the University of Michigan at Ann Arbor, and back at Oxford, where he was for seven years a Tutorial Fellow of University College.

He has held a position at Birkbeck University of London since 2005. He works mainly in philosophy of language and logic, and in the history of analytic philosophy (Frege) with forays into metaphysics and the philosophy of mathematics.

# *Truth*

# MICHAEL DUMMETT

# TRUTH

## MICHAEL DUMMETT

FREGE held that truth and falsity are the references of sentences. Sentences cannot stand for propositions (what Frege calls 'thoughts'), since the reference of a complex expression depends only on the reference of its parts; whereas if we substitute for a singular term occurring in a sentence another singular term with the same reference but a different sense, the sense of the whole sentence, *i.e.* the thought which it expresses, changes. The only thing which it appears *must* in these circumstances remain unchanged is the truth-value of the sentence. The expressions "is true" and "is false" look like predicates applying to propositions, and one might suppose that truth and falsity were properties of propositions; but it now appears that the relation between a proposition and its truth-value is not like that between a table and its shape, but rather like that between the sense of a definite description and the actual object for which it stands.

To the objection that there are non-truth-functional occurrences of sentences as parts of complex sentences, *e.g.*, clauses in indirect speech, Frege replies that in such contexts we must take ordinary singular terms as standing, not for their customary reference, but for their sense, and hence we may say that in such a context, and only then, a sentence stands for the proposition it usually expresses.

If someone asks, "But what kind of entities are these truth-values supposed to be?" we may reply that there is no more difficulty in seeing what the truth-value of a sentence may be than there is in seeing what the direction of a line may be; we have been told when two sentences have the same truth-value – when they are materially equivalent – just as we know when two lines have the same direction – when they are parallel. Nor need we waste time on the objection raised by Max Black that on Frege's theory certain sentences become meaningful which we should not normally regard as such, *e.g.,* "If oysters are inedible, then the False". If sentences stand for truth-values, but there are also expressions standing for truth-values which are not sentences, then the objection to allowing expressions of the latter kind to stand wherever sentences can stand and *vice versa* is grammatical, not logical. We often use the word "thing" to provide a noun where grammar demands one and we have only an adjective, e.g., in "That was a disgraceful thing to do"; and we could introduce a verb, say "trues", to fulfil the purely grammatical function of converting a noun standing for a truth-value into a sentence standing for the same truth-value. It may be said that

Frege has proved that a sentence does not ordinarily stand for a proposition, and has given a plausible argument that *if* sentences have references, they stand for truth-values, but that he has done nothing to show that sentences do have references at all. This is incorrect; Frege's demonstration that the notions of a concept (property) and a relation can be explained as special cases of the notion of a function provides a plausible argument for saying that sentences have a reference.

What *is* questionable is Frege's use of the words "truth" and "falsity" as names of the references of sentences; for by using these words rather than invented words of his own he gives the impression that by taking sentences to have a reference, with material equivalence as the criterion of identity, he has given an account of the notions of truth and falsity which we are accustomed to employ. Let us compare truth and falsity with the winning and losing of a board game. For a particular game we may imagine first formulating the rules by specifying the initial position and the permissible moves; the game comes to an end when there is no permissible move. We may then distinguish between two (or three) kinds of final position, which we call "Win" (meaning that the player to make the first move wins), "Lose" (similarly) and possibly "Draw". Unless we tacitly appeal to the usual meanings of the words "win", "lose" and "draw", this description leaves out one vital point – that it is the object of a player to win. It is part of the concept of winning a game that a player plays to win, and this part of the concept is not conveyed by a classification of the end positions into winning ones and losing ones. We can imagine a variant of chess in which it is the object of each player to be checkmated, and this would be an entirely different game; but the formal description we imagined would coincide with the formal description of chess. The whole theory of chess could be formulated with reference only to the formal description; but which theorems of this theory interested us would depend upon whether we wished to play chess or the variant game. Likewise, it is part of the concept of truth that we aim at making true statements; and Frege's theory of truth and falsity as the references of sentences leaves this feature of the concept of truth quite out of account. Frege indeed tried to bring it in afterwards, in his theory of assertion – but too late; for the sense of the sentence is not given in advance of our going in for the activity of asserting, since otherwise there could be people who expressed the same thoughts but went in instead for denying them.

A similar criticism applies to many accounts of truth and falsity or of the meanings of certain sentences in terms of truth and falsity. We cannot in general suppose that we give a proper account of a concept by describing those circumstances in which we do, and those in which we do not, make use of the relevant word, by describing the *usage* of that word; we must also give an account of the *point* of the concept, explain what we use the word *for*. Classifications do not exist in the void, but are connected always with some interest which we have, so that to assign

something to one class or another will have consequences connected with this interest. A clear example is the problem of justifying a form of argument, deductive or inductive. Classification of arguments into (deductively or inductively) valid and invalid ones is not a game played merely for its own sake, although it *could* be taught without reference to any purpose or interest, say as a school exercise. Hence there is really a problem of showing that the criteria we employ for recognising valid arguments do in fact serve the purpose we intend them to serve: the problem is not to be dismissed – as it has long been fashionable to do – by saying that we use the criteria we use.

We cannot assume that a classification effected by means of a predicate in use in a language will always have just *one* point. It may be that the classification of statements into true ones, false ones, and, perhaps, those that are neither true nor false, has one principal point, but that other subsidiary ends are served by it which make the use of the words "true" and "false" more complex than it would otherwise be. At one time it was usual to say that we do not call ethical statements 'true ' or 'false', and from this many consequences for ethics were held to flow. But the question is not whether these words are in practice applied to ethical statements, but whether, if they were so applied, the point of doing so would be the same as the point of applying them to statements of other kinds, and, if not, in what ways it would be different. Again, to be told that we say of a statement containing a singular term which lacks reference that it is neither true nor false is so far only to be informed of a point of usage; no philosophical consequences can yet be drawn. Rather, we need to ask whether describing such a statement as neither true nor false accords better with the general point of classifying statements as true or false than to describe it as false. Suppose that we learn that in a particular language such statements are described as 'false': how are we to tell whether this shows that they use such statements differently from ourselves or merely that "false" is not an exact translation of their word? To say that we use singular statements in such a way that they are neither true nor false when the subject has no reference is meant to characterise our use of singular statements; hence it ought to be possible to describe when in a language not containing words for "true" and "false" singular statements would be used in the same way as we use them, and when they would be used so as to be false when the subject had no reference. Until we have an account of the general point of the classification into true and false we do not know what interest attaches to saying of certain statements that they are neither true nor false; and until we have an account of how the truth-conditions of a statement determine its meaning the description of the meaning by stating the truth-conditions is valueless.

A popular account of the meaning of the word "true", also deriving from Frege, is that [it is true that P] has the same sense as the sentence P. If we then ask why it is any use to have the word "true" in the language,

the answer is that we often refer to propositions indirectly, *i.e.,* without expressing them, as when we say "Goldbach's conjecture" or "what the witness said ". We also generalise about propositions without referring to any particular one, *e.g.*, in "Everything he says is true". This explanation cannot rank as a definition in the strict sense, since it permits elimination of "is true" only when it occurs attached to a "that"-clause, and not when attached to any other expression standing for a proposition or to a variable; but, since every proposition can be expressed by a sentence, this does not refute its claim to be considered as determining uniquely the sense of "is true". It might be compared with the recursive definition of "+", which enables us to eliminate the sign "+" only when it occurs in front of a numeral, and not when it occurs in front of any other expression for a number or in front of a variable; yet there is a clear mathematical sense in which it specifies uniquely what operation "+" is to signify. Similarly, our explanation of "is true" determines uniquely the sense, or at least the application, of this predicate: for any given proposition there is a sentence expressing that proposition, and that sentence states the conditions under which the proposition is true.

If, as Frege thought, there exist sentences which express propositions but are neither true nor false, then this explanation appears incorrect. Suppose that P contains a singular term which has a sense but no reference: then, according to Frege, P expresses a proposition which has no truth-value. This proposition is therefore not true, and hence the statement [It is true that P] will be *false*. P will therefore not have the same sense as [It is true that P], since the latter is false while the former is not. It is not possible to plead that [It is true that P] is itself neither true nor false when the singular term occurring in P lacks a reference, since the *oratio obliqua* clause [that P] stands for the proposition expressed by P, and it is admitted that P does have a sense and express a proposition; the singular term occurring in P has in [lt is true that P] its indirect reference, namely its sense, and we assumed that it did have a sense. In general, it will always be inconsistent to maintain the truth of every instance of "It is true that p if and only if p" while allowing that there is a type of sentence which under certain conditions is neither true nor false. It would be possible to evade this objection by claiming that the "that"-clause in a sentence beginning "It is true that" is not an instance of *oratio obliqua*; that the word "that" here serves the purely grammatical function of transforming a sentence into a noun-clause without altering either its sense or its reference. We should then have to take phrases like "Goldbach's conjecture" and "what the witness said" as standing not for propositions but for truth-values. The expression "is true" would then be exactly like the verb "trues" which we imagined earlier; it would simply convert a noun-phrase standing for a truth-value into a sentence without altering its sense or its reference. It might be objected that this variant of Frege's account tallies badly with his saying that it is the *thought* (proposition) which is what is true or false; but we

can express this point of Frege's by saying that it is the *thought*, rather than the *sentence*, which primarily stands for a truth-value. A stronger objection to the variant account is that it leans heavily on the theory of truth-values as references of sentences, while the original version depends only on the more plausible view that clauses in indirect speech stand for propositions. In any case, if there are meaningful sentences which say nothing which is true or false, then there must be *a* use of the word "true" which applies to propositions; for if we say [lt is neither true nor false that P], the clause [that P] must here be in *oratio obliqua*, otherwise the whole sentence would lack a truth-value.

Even if we do not wish to say of certain statements that they are neither true nor false, this account cannot give the whole meaning of the word "true". If we are to give an explanation of the word "false" parallel to our explanation of "true" we shall have to say that [It is false that P] has the same sense as the negation of P. In logical symbolism there exists a sign which, put in front of a sentence, forms the negation of that sentence; but in natural languages we do not have such a sign. We have to think to realise that the negation of "No-one is here" is not "No-one is not here" but "Someone is here"; there is no one rule for forming the negation of a given sentence. Now according to what principle do we recognise one sentence as the negation of another? It is natural to answer: The negation of a sentence P is that sentence which is true if and only if P is false and false if and only if P is true. But this explanation is ruled out if we want to use the notion of the negation of a sentence in order to explain the sense of the word "false". It would not solve the difficulty if we did have a general sign of negation analogous to the logical symbol, for the question would then be: How in general do we determine the sense of the negation, given the sense of the original sentence?

We encounter the same difficulty over the connective "or". We can give an account of the meaning of "and" by saying that we are in a position to assert [P and Q] when and only when we are in a position to assert P and in a position to assert Q. (This is not circular: one could train a dog to bark only when a bell rang *and* a light shone without presupposing that it possessed the concept of conjunction.) But, if we accept a two-valued logic, we cannot give a similar explanation of the meaning of "or". We often assert [P or Q] when we are not either in a position to assert P or in a position to assert Q. I use the word "we" here, meaning mankind, advisedly. If the history master gives the schoolboy a hint, saying, "It was either James I or Charles I who was beheaded", then the schoolboy is in a position to assert, "Either James I or Charles I was beheaded" without (perhaps) being in a position to assert either limb of the disjunction; but it is not this sort of case which causes the difficulty. The ultimate source of the schoolboy's knowledge derives from something which justifies the assertion that Charles I was beheaded; and this is all that would be required for the proposed

explanation of the word "or" to be adequate. Likewise, the explanation is not impugned by cases like that in which I remember that I was talking either to Jean or to Alice, but cannot remember which. My knowledge that I was talking either to Jean or to Alice derives ultimately from the knowledge that I had at the time that I was talking to (say) Jean; the fact that the incomplete knowledge is all that survives is beside the point. Rather, the difficulty arises because we often make statements of the form [P or Q] when the ultimate evidence for making them, in the sense indicated, is neither evidence for the truth of P nor evidence for the truth of Q. The most striking instance of this is the fact that we are prepared to assert any statement of the form [P or not P], even though we may have no evidence either for the truth of P or for the truth of [Not P].

In order to justify asserting [P or not P], we appeal to the truth-table explanation of the meaning of "or". But if the whole explanation of the meanings of "true" and "false" is given by "It is true that p if and only if p" and "It is false that p if and only if not p", this appeal fails. The truth-table tells us, *e.g.,* that from P we may infer [P or Q] (in particular, [P or not P]); but that much we already knew from the explanation of "or" which we have rejected as insufficient. The truth-table does not show us that we are entitled to assert [P or not P] in every possible case, since this is to assume that every statement is either true or false; but, if our explanation of "true" and "false" is all the explanation that can be given, to say that every statement is either true or false is just to say that we are always justified in saying [P or not P].

We naturally think of truth-tables as giving the explanation of the sense which we attach to the sign of negation and to the connectives, an explanation which will show that we are justified in regarding certain forms of statement as logically true. It now appears that if we accept the redundancy theory of "true" and "false" – the theory that our explanation gives the whole meaning of these words – the truth-table explanation is quite unsatisfactory. More generally, we must abandon the idea which we naturally have that the notions of truth and falsity play an essential role in any account either of the meaning of statements in general or of the meaning of a particular statement. The conception pervades the thought of Frege that the general form of explanation of the sense of a statement consists in laying down the conditions under which it is true and those under which it is false (or better: saying that it is false under all other conditions); this same conception is expressed in the *Tractatus* in the words, "In order to be able to say that 'p' is true (or false), must have determined under what conditions I call 'p' true, and this is how I determine the sense of the sentence" (4.063). But in order that someone should gain from the explanation that P is true in such-and-such circumstances an understanding of the sense of P, he must already know what it means to say of P that it is true. If when he enquires into this he is told that the only explanation is that to say that P

is true is the same as to assert P, it will follow that in order to understand what is meant by saying that P is true, he must already know the sense of asserting P, which was precisely what was supposed to be being explained to him.

We thus have either to supplement the redundancy theory or to give up many of our preconceptions about truth and falsity. It has become a commonplace to say that there cannot be a criterion of truth. The argument is that we determine the sense of a sentence by laying down the conditions under which it is true, so that we could not first know the sense of a sentence and then apply some criterion to decide in what circumstances it was true. In the same sense there could not be a criterion for what constitutes the winning of a game, since learning what constitutes winning it is an essential part of learning what the game is. This does not mean that there may not be in any sense a theory of truth. For a particular bounded language, if it is free of ambiguity and inconsistency, it must be possible to characterise the true sentences of the language; somewhat as, for a given game, we can say which moves are winning moves. (A language is bounded if we may not introduce into it new words or new senses for old words.) Such a characterisation would be recursive, defining truth first for the simplest possible sentences, and then for sentences built out of others by the logical operations employed in the language; this is what is done for formalised languages by a truth-definition. The redundancy theory gives the general form of such a truth-definition, though in particular cases more informative definitions might be given.

Now we have seen that to say for each particular game what winning it consists in is not to give a satisfactory account of the concept of winning a game. What makes us use the same term "winning" for each of these various activities is that the point of every game is that each player tries to do what for that game constitutes winning; *i.e.*, what constitutes winning always plays the same part in determining what playing the game consists in. Similarly, what the truth of a statement consists in always plays the same role in determining the sense of that statement, and a theory of truth must be possible in the sense of an account of what that role is. I shall not now attempt such an account; I claim, however, that such an account would justify the following. A statement, so long as it is not ambiguous or vague, divides all possible states of affairs into just *two* classes. For a given state of affairs, either the statement is used in such a way that a man who asserted it but envisaged that state of affairs as a possibility would be held to have spoken misleadingly, or the assertion of the statement would not be taken as expressing the speaker's exclusion of that possibility. If a state of affairs of the first kind obtains, the statement is false; if all actual states of affairs are of the second kind, it is true. It is thus *prima facie* senseless to say of any statement that in such-and-such a state of affairs it would be neither true nor false.

The sense of a statement is determined by knowing in what circumstances it is true and in what false. Likewise the sense of a command is determined by knowing what constitutes obedience to it and what disobedience; and the sense of a bet by knowing when the bet is won and when it is lost. Now there may be a gap between the winning of a bet and the losing of it, as with a conditional bet; can there be a similar gap between obedience and disobedience to a command, or between the truth and falsity of a statement? There is a distinction between a conditional bet and a bet on the truth of a material conditional; if the antecedent is unfulfilled, in the first case the bet is off – it is just as if no bet had been made – but in the second case the bet is won. A conditional command where the antecedent is in the power of the person given the order (*e.g.,* a mother says to a child, "If you go out, wear your coat") is always like a bet on the material conditional; it is equivalent to the command to ensure the truth of the material conditional, *viz.*, "Do not go out without your coat". We cannot say that if the child does not go out, it is just as if no command had been given, since it may be that, unable to find his coat, he stayed in in order to comply with the command.

Can a distinction parallel to that for bets be drawn for conditional commands where the antecedent is not in the person's power? I contend that the distinction which looks as if it could be drawn is in fact void of significance. There are two distinct kinds of consequence of making a bet, winning it and losing; to determine what is to involve one of these is not yet to determine completely what is to involve the other. But there is only one kind of consequence of giving a command, namely that, provided one had the right to give it in the first place, one acquires a right to punish or at least reprobate disobedience. It might be thought that punishment and reward were distinct consequences of a command in the same sense that paying money and receiving it are distinct consequences of a bet; but this does not tally with the role of commands in our society. The right to a reward is not taken to be an automatic consequence of obedience to a command, as the *right* to reproach is an automatic consequence of disobedience; if a reward is given, this is an act of grace, just as it is an act of grace if the punishment or reproach is withheld. Moreover, any action deliberately taken in order to comply with the command (to avoid disobedience to it) has the same claim to be rewarded as any other; hence to determine what constitutes disobedience to the command is thereby to determine what sort of behaviour might be rewarded, without the need for any further decision. If the child stays in because he cannot find his coat, this behaviour is as meritorious as if he goes out remembering to wear it; and if he forgets all about the order, but wears his coat for some other reason, this behaviour no more deserves commendation than if he chooses, for selfish reasons, to remain indoors. Where the antecedent is not in the person's power, it is indeed possible to regard the conditional command as analogous to the conditional bet; but since obedience to a command has no consequence

of its own other than that of avoiding the punishment due for disobedience, there is not for such commands any significant distinction parallel to that between conditional bets and bets about a material conditional. If we regarded obedience to a command as giving a right to a reward, we could then introduce such a distinction for commands whose antecedent was in the person's power. Thus the mother might use the form, "If you go out, wear your coat", as involving that if the child went out with his coat he would be rewarded, if he went out without it he would be punished, and if he stayed indoors – even in order to comply with the command – he would be neither punished nor rewarded; while the form, "Do not go out without your coat ", would involve his being rewarded if he stayed indoors.

Statements are like commands (as we use them) and not like bets; the making of a statement has, as it were, only one kind of consequence. To see this, let us imagine a language which contains conditional statements but has no counterfactual form (counterfactuals would introduce irrelevant complications). Two alternative accounts are suggested of the way in which conditionals are used in this language: one, that they are used to make statements conditionally; the other, that they represent the material conditional. On the first interpretation, a conditional statement is like a conditional bet: if the antecedent is fulfilled, then the statement is treated as if it had been an, unconditional assertion of the consequent, and is said to be true or false accordingly; if the antecedent is not fulfilled, then it is just as if no statement, true or false, had been made at all. On the second interpretation, if the antecedent is not fulfilled, then the statement is said to be true. How are we to settle which of these two accounts is the correct one? If statements are really like bets and not like commands; if there are two distinct kinds of consequence which may follow the making of a statement, those that go with calling the statement 'true' and those that go with calling it ' false ', so that there may be a gap between these two kinds of consequence; then we ought to be able to find something which decides between the two accounts as definite as the financial transaction which distinguishes a bet on the truth of the material conditional from a conditional bet. It is no use asking whether these people say that the man who has made a conditional statement whose antecedent turns out false said something true or that he said nothing true or false: they may have no words corresponding to "true" and "false"; and if they do, how could we be sure that the correspondence was exact? If their using the words "true" and "false" is to have tle slightest significance, there must be some difference in their behaviour which goes with their saying "true" or neither "true nor false" in this case.

It is evident on reflection that there is nothing in what they do which could distinguish between the two alternative accounts; the distinction between them is as empty as the analogous distinction for conditional commands whose antecedent is not in the person's power. In order to fix

the sense of an utterance, we do not need to make two separate decisions – when to say that a true statement has been made and when to say that a false statement has been made; rather, any situation in which nothing obtains which is taken as a case of its being false may be regarded as a case of its being true, just as someone who behaves so as not to disobey a command may be regarded as having obeyed it. The point becomes clearer when we look at it in the following way. If it makes sense in general to suppose that a certain form of statement is so used that in certain circumstances it is true, in others false, and in yet others nothing has been said true or false, then we can imagine that a form of conditional was used in this way (von Wright actually holds that we use conditionals in this way). If P turns out true, then [If P, then Q] is said to be true or false according as Q is true or false, while if P turns out false we say that nothing was said true or false. Let us contrast this with what Frege and Strawson say about the use in our language of statements containing a singular term. If there is an object for which the singular term stands, then the statement is true or false according as the predicate does or does not apply to that object, but if there is no such object, then we have not said anything true or false. Now do these accounts tell us the sense of sentences of these two kinds? – that is, do they tell us how these statements are used, what is *done* by making statements of these forms? Not at all, for an essential feature of their use has not yet been laid down. Someone uttering a conditional statement of the kind described may very well have no opinion as to whether the antecedent was going to turn out true or false; that is, he is not taken as having misused the statement or misled his hearers if he envisages it as a possibility that that case will arise in which he is said not to have made a statement true or false. All that he conveys by uttering the conditional statement is that he excludes the possibility that the case will arise in which he is said to have said something false, namely that antecedent is true and consequent false. With the case of a singular statement it is quite different. Here someone is definitely either misusing the form of statement or misleading his hearers if he envisages it as a possibility that that case will arise in which what he said will be said to be neither true nor false, namely that the singular term has no reference. He conveys more by making the statement than just that he excludes the possibility of its being false; he commits himself to its being true.

Are we then to say that laying down the truth-conditions for a sentence is not sufficient to determine its sense, that something further will have to be stipulated as well? Rather than say this we should abandon the notions of truth and falsity altogether. In order to characterise the sense of expressions of our two forms, only a twofold classification of possible relevant circumstances is necessary. We need to distinguish those states of affairs such that if the speaker envisaged them as possibilities he would be held to be either misusing the statement or misleading his hearers, and those of which this is not the case: and *one* way of using the words "true" and "false" would be to call states of

affairs of the former kind those in which the statement was false and the others those in which the statement was true. For our conditional statements, the distinction would be between those states of affairs in which the statement was said to be false and those in which we said that it would either be true or else neither true nor false. For singular statements, the distinction would be between those states of affairs in which we said that the statement would either be false or else neither true nor false, and those in which it was true. To grasp the sense or use of these forms of statement, the twofold classification is quite sufficient; the threefold classification with which we started is entirely beside the point. Thus, on *one* way of using the words "true" and "false", we should, instead of distinguishing between the conditional statement's being true and its being neither true nor false, have distinguished between two different ways in which it could be true; and instead of distinguishing between the singular statement's being false and its being neither true nor false, we should have distinguished between two different ways in which it could be false.

This gives us a hint at a way of explaining the rôle played by truth and falsity in determining the sense of a statement. We have not yet seen what point there may be in distinguishing between different ways in which a statement may be true or between different ways in which it may be false, or, as we might say, between degrees of truth and falsity. The point of such distinctions does not lie in anything to do with the sense of the statement itself, but has to do with the way in which it enters into complex statements. Let us imagine that in the language of which the conditional statements we considered form part there exists a sign of negation, *i.e.*, a word which, placed in front of a statement, forms another statement; I call it a sign of negation because in most cases it forms a statement which we should regard as being used as the contradictory of the original statement. Let us suppose, however, that when placed in front of a conditional statement [If P, then Q], it forms a statement which is used in the same way as the statement 'If P, then not Q'. Then if we describe the use of the conditionals by reference to a twofold classification only, *i.e.*, in the same way as we describe a material conditional, we shall be unable to give a truth-functional account of the behaviour of their sign "not". That is, we should have the tables:

| P | Q | [If P, then Q] | [Not: if P, then Q] |
|---|---|---|---|
| T | T | T | F |
| T | F | F | T |
| F | T | T | T |
| F | F | T | T |

in which the truth-value of [Not: if P, then Q] is not determined by the truth-value of [If P, then Q]. If, on the other hand, we revert to our

original threefold classification, marking the case in which we said that no statement true or false had been made by "X ", then we have the tables:

| P | Q | [If P, then Q] | [Not: if P, then Q] |
|---|---|---|---|
| T | T | T | F |
| T | F | F | T |
| F | T | X | X |
| F | F | X | X |

which can be quite satisfactorily accounted for by giving the table for "not":

| R | [Not R] |
|---|---|
| T | F |
| X | X |
| F | T |

(I have assumed that the statements P and Q take only the values T and F.) It now becomes quite natural to think of "T" as representing "true", "F" "false" and "X" "neither true nor false". Then we can say that their symbol "not" really is a sign of negation, since [Not P] is true when and only when P is false and false when and only when P is true. We must not forget, however, that the justification for distinguishing between the cases in which a conditional was said to have the value T and the cases in which it was said to have the value X was simply the possibility, created by this distinction, of treating "not" truth-functionally. In the same way if we have in a language an expression which normally functions as a sign of negation, but the effect of prefacing a singular statement with this expression is to produce a statement whose utterance still commits the speaker to there being an object for which the singular term stands, it is very natural to distinguish between two kinds of falsity a singular statement may have: that when the singular term has a reference, but the predicate does not apply to it, and that when the singular term lacks a reference. Let us represent the case in which the singular term has no reference by the symbol "Y", and let us suppose S to be a singular statement. Then we have the table:

| S | [Not S] |
|---|---|
| T | F |
| Y | Y |
| F | T |

Here again it is natural to think of "T" as representing "true", "F" "false" and "Y" "neither true nor false".

There is no necessity to use the words "true" and "false" as suggested above, so that we have to interpret X as a kind of truth and Y as a kind of falsity. Logicians who study many-valued logics have a term which can be employed here: they would say that T and X are 'designated' truth-values and F and Y 'undesignated' ones. (In a many-valued logic those formulae are considered valid which have a designated value for every assignment of values to their sentence-letters.) The points to observe are just these. (i) The sense of a sentence is determined wholly by knowing the case in which it has a designated value and the cases in which it has an undesignated one. (ii) Finer distinctions between different designated values or different undesignated ones, however naturally they come to us, are justified only if they are needed in order to give a truth-functional account of the formation of complex statements by means of operators. (iii) In *most* philosophical discussions of truth and falsity, what we really have in mind is the distinction between a designated and an undesignated value, and hence choosing the names "truth" and "falsity" for particular designated and undesignated values respectively will only obscure the issue. (iv) Saying that in certain circumstances a statement is neither true nor false does not determine whether the statement is in that case to count as having an undesignated or a designated value, *i.e.*, whether someone who asserts the statement is or is not taken as excluding the possibility that that case obtains.

Baffled by the attempt to describe in general the relation between language and reality, we have nowadays abandoned the correspondence theory of truth, and justify our doing so on the score that it was an attempt to state a *criterion* of truth in the sense in which this cannot be done. Nevertheless, the correspondence theory expresses one important feature of the concept of truth which is not expressed by the law "It is true that p if and only if p" and which we have so far left quite out of account: that a statement is true only if there is something in the world in virtue of which it is true. Although we no longer accept the correspondence theory, we remain realists *au fond*; we retain in our thinking a fundamentally realist conception of truth. Realism consists in the belief that for any statement there must be something *in virtue of which* either it or its negation is true: it is only on the basis of this belief that we can justify the idea that truth and falsity play an essential role in the notion of the meaning of a statement, that the general form of an explanation of meaning is a statement of the truth-conditions.

To see the importance of this feature of the concept of truth, let us envisage a dispute over the logical validity of the statement "Either Jones was brave or he was not". A imagines Jones to be a man, now dead, who never encountered danger in his life. B retorts that it could still be true that Jones was brave, namely if it is true that if Jones *had* encountered danger, he would have acted bravely. A agrees with this, but still maintains that it does not need to be the case that either "Jones was brave" = "If Jones had encountered danger, he would have acted

bravely" nor "Jones was not brave" If Jones had encountered danger, he would not have acted bravely" is true. For, he argues, it might be the case that however many facts we knew of the kind which we should normally regard as grounds for asserting such counterfactual conditionals, we should still know nothing which would be a ground for asserting either. It is clear that B cannot agree that this is a possibility and yet continue to insist that all the same either "Jones was brave" or "Jones was not brave" is true; for he would then be committed to holding that a statement may be true even though there is nothing whatever such that, if we knew of it, we should count it as evidence or as a ground for the truth of the statement, and this is absurd. (It may be objected that there are assertions for which it would be out of place to ask one who made them for his evidence or grounds; but for *such* assertions the speaker must always either be in a position to make or in a position to deny them.) If B still wishes to maintain the necessity of "Either Jones was brave or he was not", he will have to hold either that there must be some fact of the sort to which we usually appeal in discussing counterfactuals which, if we knew it, would decide us in favour either of the one counterfactual or of the other; or else that there is some fact of an extraordinary kind, perhaps known only to God. In the latter case he imagines a kind of spiritual mechanism – Jones' character – which determines how he acts in each situation that arises; his acting in such-and-such a way reveals to us the state of this spiritual mechanism, which was however already in place before its observable effects were displayed in his behaviour. B would then argue thus: If Jones had encountered danger, he would either have acted bravely or have acted like a coward. Suppose he had acted bravely. This would then have shown us that he was brave; but he would *already* have been brave before his courage was revealed by his behaviour. That is, either his character included the quality of courage or it did not, and his character determines his behaviour. We know his character only indirectly, through its effects on his behaviour; but each character-trait must be *there* within him independently of whether it reveals itself to us or not.

Anyone of a sufficient degree of sophistication will reject B's belief in a spiritual mechanism; either he will be a materialist and substitute for it an equally blind belief in a physiological mechanism, or he will accept A's conclusion that "Either Jones was brave or he was not" is not logically necessary. His ground for rejecting B's argument is that if such a statement as "Jones was brave" is true, it must be true in virtue of the sort of fact we have been taught to regard as justifying us in asserting it. It cannot be true in virtue of a fact of some quite different sort of which we can have no direct knowledge, for otherwise the statement "Jones was brave" would not have the meaning that we have given it. In accepting A's position he makes a small retreat from realism; he abandons a realist view of character.

In order, then, to decide whether a realist account of truth can be given for statements of some particular kind, we have to ask whether for such a statement P it must be the case that if we knew sufficiently many facts of the kind we normally treat as justifying us in asserting P, we should be in a position either to assert P or to assert [Not P]: if so, then it can truly be said that there must either be something in virtue of which P is true or something in virtue of which it is false. It is easy to overlook the force of the phrase "sufficiently many". Consider the statement "A city will never be built on this spot ". Even if we have an oracle which can answer every question of the kind, "Will there be a city here in 1990?" "In 2100?" *etc.*, we might never be in a position either to declare the statement true or to declare it false. Someone may say: That is only because you are assuming the knowledge of only finitely many answers of the oracle; but if you knew the oracle's answers to all these questions, you would be able to decide the truth-value of the statement. But what would it mean to know infinitely many facts? It could mean that the oracle gave a direct answer "No" to the question, "Will a city ever be built here?": but to assume this is just like B's assumption of the existence of a hidden spiritual mechanism. It might mean that we had an argument to show the falsity of FA city will be built here in the year N irrespective of the value of N, *e.g.*, if 'here' is the North Pole: but no-one would suggest that it must be the case that either the oracle will give an affirmative answer to some question of the form "Will there be a city here in the year .... ?" or we can find a general argument for a negative answer. Finally, it could mean that we were *able* to answer every question of the form, "Will there be a city here in the year .... ? ": but having infinite knowledge in *this* sense will place us in no better position than when we had the oracle.

We thus arrive at the following position. We are entitled to say that a statement P must be either true or false, that there must be something in virtue of which either it is true or it is false, only when P is a statement of such a kind that we could in a finite time bring ourselves into a position in which we were justified either in asserting or in denying P; that is, when P is an effectively decidable statement. This limitation is not trivial: there is an immense range of statements which, like "Jones was brave", are concealed conditionals, or which, like "A city will never be built here", contain – explicitly or implicitly – an unlimited generality, and which therefore fail the test.

What I have done here is to transfer to ordinary statements what the intuitionists say about mathematical statements. The sense of *e.g.,* the existential quantifier is determined by considering what sort of fact makes an existential statement true, and this means: the sort of fact which we have been taught to regard as justifying us in asserting an existential statement. What would make the statement that there exists an odd perfect number true would be some particular number's being both odd and perfect; hence the assertion of the existential statement

must be taken as a claim to be able to assert some one of the singular statements. We are thus justified in asserting that there is a number with a certain property only if we have a method for finding a particular number with that property. Likewise, the sense of a universal statement is given by the sort of consideration we regard as justifying us in asserting it: namely we can assert that every number has a certain property if we have a general method for showing, for any arbitrary number, that it has that property. Now what if someone insists that either the statement "There is an odd perfect number" is true, or else every perfect number is even? He is justified if he knows of a procedure which will lead him in a finite time either to the determination of a particular odd perfect number or to a general proof that a number assumed to be perfect is even. But if he knows of no such procedure, then he is trying to attach to the statement "Every perfect number is even" a meaning which lies *beyond* that provided by the training we are given in the use of universal statements; he wants to say, as B said of "Jones was brave", that its truth may lie in a region directly accessible only to God, which human beings can never survey.

We learn the sense of the logical operators by being trained to *use* statements containing them, *i.e.*, to assert such statements under certain conditions. Thus we learn to assert [P and Q] when we can assert P and can assert Q; to assert [P or Q] when we can assert P or can assert Q; to assert [For some n, F(n)] when we can assert [F(O)] or can assert [F(1)] or… We learn to assert [For every n, F(n)] when we can assert [F(O)] and [F(1)] and…. ; and to say that we can assert all of these means that we have a general method for establishing [F(x)] irrespective of the value of x. Here we have abandoned altogether the attempt to explain the meaning of a statement by laying down its truth-conditions. *We no longer explain the sense of a statement by stipulating its truth-value in terms of the truth-values of its constituents, but by stipulating when it may be asserted in terms of the conditions under which its constituents may be asserted*. The justification for this change is that this is how we in fact learn to use these statements: furthermore, the notions of truth and falsity cannot be satisfactorily explained so as to form a basis for an account of meaning once we leave the realm of effectively decidable statements. One result of this shift in our account of meaning is that, unless we are dealing only with effectively decidable statements, certain formulae which appeared in the two-valued logic to be logical laws no longer rank as such, in particular the law of excluded middle: this is rejected, not on the ground that there is a middle truth-value, but because meaning, and hence validity, is no longer to be explained in terms of truth-values.

Intuitionists speak of mathematics in a highly anti-realist (anti-platonist) way: for them it is *we* who construct mathematics; it is not already *there* waiting for us to discover. An extreme form of such constructivism is found in Wittgenstein's *Remarks on the Foundations of*

*Mathematics.* This makes it appear as though the intuitionist rejection of an account of the meaning of mathematical statements in terms of truth and falsity could not be generalised for other regions of discourse, since even if there is no independent mathematical reality answering to our mathematical statements, there is an independent reality answering to statements of other kinds. On the other hand the exposition of intuitionism I have just given was not based on a rejection of the Fregean notion of a mathematical reality waiting to be discovered, but only on considerations about meaning. Now certainly someone who accepts the intuitionist standpoint in mathematics will not be inclined to adopt the platonist picture. Must he then go to the other extreme, and have the picture of our creating mathematics as we go along? To adopt this picture involves thinking with Wittgenstein that we are *free* in mathematics at every point; no step we take has been forced on us by a necessity external to us, but has been freely chosen. This picture is not the only alternative. If we think that mathematical results are in some sense imposed on us from without, we could have instead the picture of a mathematical reality not already in existence but as it were coming into being as we probe. Our investigations bring into existence what was not there before, but what they bring into existence is not of our own making.

Whether this picture is right or wrong for mathematics, it is available for other regions of reality as an alternative to the realist conception of the world. This shows how it is possible to hold that the intuitionist substitution of an account of the *use* of a statement for an account of its truth-conditions as the general form of explanation of meaning should be applied to all realms of discourse without thinking that we create the world; we can abandon realism without falling into subjective idealism. This substitution does not of course involve dropping the words "true" and "false", since for most ordinary contexts the account of these words embodied in the laws "It is true that p if and only if p" and "It is false that p if and only if not p" is quite sufficient: but it means facing the consequences of admitting that this is the *whole* explanation of the sense of these words, and this involves dethroning truth and falsity from their central place in philosophy and in particular in the theory of meaning. Of course the doctrine that meaning is to be explained in terms of use is the cardinal doctrine of the later Wittgenstein; but I do not think the point of this doctrine has so far been generally understood.

*Michael Dummett's 'Truth'*
# IAN RUMFITT

# MICHAEL DUMMETT'S 'TRUTH'

## IAN RUMFITT

MICHAEL Dummett read his paper 'Truth' to the Aristotelian Society on Monday 16 February 1959 during a critical period in his philosophical development. The previous year, Dummett had submitted to Oxford University Press a book, *The Law of Excluded Middle*; this was his first sustained attempt to address the problem of justifying basic logical laws that was to pre-occupy him for the rest of his life. As he recalled nearly twenty years later in the Preface to *Truth and Other Enigmas* (Dummett 1978, pp.xix-xx), the Press accepted the book on the advice of J.L. Austin, who was then one of its delegates; Austin, however, had reservations about Dummett's literary style and required substantial changes. 'As I engaged in the laborious process of trying to comply', Dummett wrote, 'I became more and more dissatisfied with the content of the book, and never resubmitted it. In a sense, I have been trying to rewrite the book ever since' (*ibid*.).[1] 'Truth' is a product of this process of reconsidering his early, markedly Wittgensteinian views.

The paper is rich. Indeed, it stands alongside Quine's 'Truth by convention' (1936) as an example of an early essay in which many of a major philosopher's central concerns and contentions are adumbrated and prefigured. The end of the article (see esp. 161[2]) already recommends the shift, from a theory of meaning based on truth-conditions to one based on assertability-conditions, that Dummett was to explore more fully in his writings of the 1970s; as he recognizes (*ibid*.), this shift involves demoting Excluded Middle from its status as a logical truth. In the spirit of this symposium, however, I will focus these introductory remarks not on the theory of meaning, or on the justification of logical laws, but on what Dummett says about truth.

---

[1] After Dummett's death in December 2011, his family and literary executors (of whom I am one) searched the house that he and his wife Ann had shared for fifty-five years in the hope of finding a copy of the typescript of *The Law of Excluded Middle*. Although the older Dummett said he would have been ashamed of this early book, had it been published (see Dummett 2007, 15), this judgement was not based on any recent re-reading of the work, and may well have been too harsh. Unfortunately, the typescript was nowhere to be found.

[2] Unadorned page references are to the original printing of 'Truth' in the *Proceedings of the Aristotelian Society*.

The paper is shaped by an extended comparison between the notions of truth and falsity and those of winning and losing a game. We could, Dummett remarks, classify all the possible final positions in a game of chess into those in which the first player to move wins, or loses, or draws (142). We could make the corresponding classification for the game of draughts. These classifications, though, do not capture what it is to win at a game, nor what winning at chess and winning at draughts have in common. There is a similar inadequacy, Dummett thinks, in extant philosophical theories of truth. We could in principle construct a theory which yields, for each statement that could be made in English, the condition which must be satisfied for it to be true. We could construct a corresponding theory for German. Such theories, though, reveal little about the concept of truth, because they fail to articulate what it is for a statement to be true, or what true statements made in different languages have in common. The analogy with games also indicates what is missing. According to Dummett, it is part of the concept of winning a game—any game—that a player plays to win (142). Similarly, 'it is part of the concept of truth that we aim at making true statements' (143). But these formulae do not supply all of what we need, and in any event they are at best only approximately correct. Few of those who played chess against Josef Stalin played to win, although no doubt all his opponents pretended to do so. Similarly, liars do not aim at making true statements, although they differ from jokers in representing themselves as making true statements.

Dummett holds that the general concept of winning must be invoked in any full account of what it is to play a particular game. Indeed, 'what constitutes winning always plays the *same* part in determining what playing the game consists in. Similarly, what the truth of a statement consists in always plays the same rôle in determining the sense of that statement, and a theory of truth must be possible in the sense of an account of what that rôle is' (149, emphasis added). Disappointingly, Dummett does not essay such an account in the present paper, but knowing how he described that rôle in his later writings is, I think, necessary to understanding one of his main theses here. An account of that rôle, he claims in 'Truth',

> would justify the following. A statement, so long as it is not ambiguous or vague, divides all possible states of affairs into just *two* classes. For a given state of affairs, either the statement is used in such a way that a man who asserted it but envisaged that state of affairs as a possibility would be held to have spoken misleadingly, or the assertion of the statement would not be taken as expressing the speaker's exclusion of that possibility. If a state of affairs of the first kind obtains, the

statement is false; if all actual states of affairs are of the second kind, it is true. It is thus *prima facie* senseless to say of any statement that in such-and-such a state of affairs it would be neither true nor false (149-150).

Dummett relies here on what we may call an *exclusionary* account of content: a statement's content is given by the possibilities that it excludes. In 'Truth', he gives no reason for accepting this account, but elsewhere he does. For Dummett, a statement's content is always the content of a potential assertion of it, and in his article 'What is a theory of meaning? (II)' of 1976, he proposes an account of assertion which supports his exclusionary theory of content. 'An assertion', he observes in the later article, 'is not, normally, like an answer in a quiz programme; the speaker gets no prize for being right. It is, primarily, a guide to action on the part of the hearers (an interior judgment being a guide to action on the part of the thinker); a guide which operates by inducing in them certain expectations. And the content of an expectation is determined by what will surprise us; that is, by what it is that is *not* in accord with the expectation rather than by what corroborates it' (Dummett 1976, 124). The content of an assertion, then, is given by the states of affairs that one who understands and accepts it is thereby entitled to set aside or exclude—just as 'Truth' proposes. Moreover, the states of affairs whose exclusion determines content must be practically relevant: they must be such that their obtaining, or not, potentially bears upon whether a course of action is rational for a hearer.

How plausible is the account of truth and falsity that Dummett erects on his exclusionary theory of content? It fits *some* of his examples rather well. Consider the statement *S*, 'A city will never be built on this spot', as uttered in place $\pi$ on day $d$ (159). Statement *S* excludes every member of the following set of possible states of affairs:{a city stands at $\pi$ at the end of day $d+1$, a city stands at $\pi$ at the end of day $d+2$,...}. (In the spirit of the proverb about Rome, I assume it takes at least a day to build a city.) Moreover, all the practically relevant states of affairs that are excluded by *S* belong to this set. While *S* may be said to exclude the possibility that a city should one day stand at $\pi$, this possibility does not constitute a practically relevant state of affairs. In the words of 'What is a theory of meaning? (II)', such a possibility 'has no substance', for the expectation that it obtains cannot be disappointed, and hence cannot guide rational action. Now according to Dummett, a statement is true when (and presumably only when) no excluded state of affairs obtains, and is false when (and presumably only when) an excluded state of affairs does obtain. Applying these accounts to *S*, then, we reach

> True $(S) \leftrightarrow \forall n$ (no city stands at $\pi$ at the end of day $d + n$)

and

> False $(S) \leftrightarrow \exists n$ (a city stands at $\pi$ at the end of day $d + n$).

These specifications of the truth- and falsity-conditions of $S$ seem correct: Dummett's requirement of practical relevance does not stop his theory from delivering the desired results in this case. It is worth considering, though, whether the account generates the intuitively correct results in other cases too.

However that may be, there is a deep tension between the broadly pragmatist account of truth that generates these truth-conditions and the verificationist theses that Dummett advances later in the essay. There is, he writes, an 'important feature of the concept of truth which…we have so far left quite of account: that a statement is true only if there is something in the world *in virtue of which* it is true' (157). The things in virtue of which a statement is true, he explains, are 'the sort of fact we have been taught to regard as justifying us in asserting it' (159). Let us call such a fact a *ground* of the statement. Then Dummett is making the following claim:

> $(T^G)$     $\forall u$(True $(u) \rightarrow$ a ground for $u$ obtains).

He also maintains the corresponding thesis about falsehood. Let us say that a fact is an *anti-ground* of a statement if its obtaining justifies us in denying the statement. Then we have

> $(F^G)$     $\forall u$(False $(u) \rightarrow$ an anti-ground for $u$ obtains).

$(T^G)$ and $(F^G)$ together entail

> $(Biv^G)$     $\forall u$((True $(u) \lor$ False $(u)) \rightarrow$ (a ground for $u$ obtains $\lor$ an anti-ground for $u$ obtains).

Dummett's pragmatist accounts of truth and falsity explain why he holds that 'it is *prima facie* senseless to say of any statement that in such-and-such a state of affairs it would be neither true nor false'. In particular, then, it is *prima facie* senseless to say that $S$ is neither true nor false. Dummett maintains, however, that we cannot assert that $S$ is

bivalent, and we can now understand why he says this.[3] If we were to assert that $S$ is either true or false then it would follow, by ($Biv^G$), that either a ground or an anti-ground for $S$ obtains. In many circumstances, though, this condition will not be met:

> We are entitled to say that a statement $P$ must be either true or false, that there must be something in virtue of which either it is true or it is false, only when $P$ is a statement of such a kind that we could in a finite time bring ourselves into a position in which we are justified either in asserting or in denying $P$; that is, when $P$ is an effectively decidable statement. This limitation is not trivial: there is an immense range of statements which, like 'Jones was brave' [said of a man who died without facing danger], are concealed conditionals, or which, like 'A city will never be built here', contain—explicitly or implicitly—an unlimited generality, and which therefore fail the test (160).

I agree with Dummett that in many circumstances we are not entitled to assert that either a ground or an anti-ground for $S$ obtains. A ground for $S$ is the sort of fact that justifies us in asserting 'A city will never be built here'. We know what sort of facts these are: the place in question is too cold, or is too hot, or lacks sufficient water, etc. Let us suppose that none of these features afflict $\pi$. The locus of $S$, in other words, is a place where a city *could* well be built. On that supposition, no ground for $S$ obtains. An anti-ground for $S$ is the sort of fact that justifies us in denying 'A city will never be built here'. Again, we know what sort of facts *these* are: plans to build a city are already laid, or population growth makes the urbanization of the relevant place inevitable, etc. Consistently with our first supposition, we may further suppose that none of these facts obtain either. In that case, we shall not be entitled to say that either a ground or an anti-ground for $S$ obtains.

But does it follow that we cannot assert that $S$ is bivalent? I do not think so. For one might instead take $S$ to be a counterexample to ($Biv^G$) and ($T^G$). Indeed, it seems to be a particularly strong form of counterexample to ($T^G$), for $S$ may be true even though a ground for $S$ *never* obtains. It might be that $\pi$ always remains a place where a city *could* be built—so that no one is ever justified in asserting 'A city will never be built here'—while, as a matter of fact, no one ever happens to build a city there—so that 'A city will never be built here' is true.

---

[3] In classical logic, 'No statement is neither true nor false' entails 'Every statement is either true or false', so Dummett's position is coherent only if we revise classical logic. For a discussion of this aspect of his argument, see my 2007.

In this regard, it is helpful to compare Dummett's *S* with Aristotle's statement, *T*, 'There will be a sea-battle tomorrow', made, let us suppose, on today, 1 January. Because the admirals have not decided whether to fight, no one today would be justified in asserting that *T* is true and no one today would be justified in asserting that *T* is false. Tomorrow, though, some people will be justified in asserting either that *T* is true or that it is false; by 2 January, either a ground or an anti-ground will have obtained. The proper moral of Dummett's example *S*, a classical logician might think, is that even this *eventual* obtaining of either a ground or an anti-ground for a statement is inessential to its bivalence. If Dummett is right to deny that *S* is neither true nor false, then it follows that we may assert that it is bivalent. We may assert that *S* is either true or false even though we know that neither a ground nor an anti-ground for *S* may ever obtain. His case brings out something interesting, then, but it is not a counterexample to Bivalence.

The tension between pragmatist and verificationist elements in the theory of meaning and in the account of truth was one that Dummett struggled to resolve. He sought, but never fully achieved, an overview of the relationship between the grounds for making an assertion and the effects it would have on a hearer who understood and accepted it. No one else, however, has attained such an overview, and few of today's philosophers even try. One reason for reading 'Truth' is that it brings this unsolved problem very starkly to our attention.

REFERENCES

Auxier, R.E. and L.E. Hahn, eds. 2007. *The Library of Living Philosophers Volume XXXI: The Philosophy of Michael Dummett*. Chicago, Ill.: Open Court.

Dummett, M.A.E. 1976. 'What is a theory of meaning? (II)'. In Evans and McDowell, eds., *Truth and Meaning: Essays in Semantics* (Oxford: Clarendon Press), pp.67-137.

—— . 1978. *Truth and Other Enigmas*. London: Duckworth.

—— . 2007. 'Intellectual autobiography'. In Auxier and Hahn, eds., 2007, pp.3-32.

Quine, W.V. 1936. 'Truth by convention'. In O.H. Lee, ed., *Philosophical Essays for A.N. Whitehead* (New York: Longmans), pp.90-124.

Rumfitt, I. 2007. 'Asserting and excluding'. In Auxier and Hahn, eds., 2007, pp.639-693.

PETER GEACH & GRAHAM OPPY

# BIOGRAPHY

Peter Thomas Geach, MA, FBA (born 29 March 1916) is a British philosopher with areas of interest in the history of philosophy, philosophical logic, and the theory of identity. Geach was educated at Balliol College, Oxford. He taught at the University of Birmingham from 1951 until 1966 when he was appointed Professor of Logic in the Department of Philosophy at the University of Leeds. Geach was given the title of Emeritus Professor of Logic on his retirement from Leeds in 1981. He has been awarded the papal cross "Pro Ecclesia et Pontifice" by the Holy See for his philosophical work.

"Truth and God" was originally published in the *Proceedings of the Aristotelian Society, Supplementary Volume LVI* (1982).

Graham Oppy is Professor of Philosophy in the School of Philosophical, Historical and International Studies at Monash University. His books include: *Ontological Arguments and Belief in God*, *Philosophical Perspectives on Infinity*, *Arguing about Gods*, *Reading Philosophy of Religion* (co-authored with Michael Scott), *The History of Western Philosophy of Religion* (co-edited with Nick Trakakis), *A Companion to Philosophy in Australia and New Zealand* (co-edited with Nick Trakakis), and *The Antipodean Philosopher* (two volumes, co-edited with Nick Trakakis). He is a Fellow of the Australian Academy of Humanities, and Chair of Council of the Australasian Association of Philosophy.

*Truth and God*

# PETER GEACH

# TRUTH AND GOD

## PETER GEACH

WHEN we consider truth in a quite particular example, there seems to be no problem about truth as such. Suppose someone gravely asks me: 'Is it true that our sovereign liege lady is deceased?' If I once understand that his last six words just mean 'the Queen is dead', then his question amounts to no more and no less than 'Is the Queen dead?'; and to answer that we need advert to no problem about truth. Death raises serious problems, both philosophically and otherwise, but not specially about truth.

However, we cannot confine our use of 'true' and 'truth' to such examples. Counsel assures the jury in a trial for motor manslaughter that what P.C.49 says on oath will be true, since he is a man of great integrity and a reliable observer. (We may here notice a difficulty for the inductivist theory of testimony, so often light-mindedly accepted since Hume's day: how can we reliably establish by induction a correlation between *being what P.C.49 says on oath* and *being true? Being true* is not, surely, on a level with *being blue* or Goodman's *being grue*.) From 'What P.C.49 says is true' we cannot extract something to the same effect and no longer involving mention of truth. It is because of such apparently uneliminable uses of 'true' that philosophers have come to construct theories of truth.

Some theories of truth turn out to be blunderings or blind alleys: the difficulty of showing this varies from case to case. By the speech-act theory of truth, 'What P.C.49 says is true' (e.g.) would be assimilated to a sentence with a verb used performatively, like 'I corroborate what P.C.49 says'; this theory hardly gets airborne; it ignores all uses of 'is true' except as the single main predicate of an assertoric sentence, although like other grammatical predicates 'is true' can occur in dozens of other sorts of context. The predicable 'is a scandalous revelation if true', for example, cannot be twisted into any form employing a performative verb in place of its constituent '(is) true'.

Another theory that may be quickly dismissed is the theory that truth is a kind of correspondence relation to a fact. The identity of facts as entities raises many problems: I am inclined myself to think that the problems are intractable, and facts consequently *entia non grata*, in Quine's phrase. Luckily we need not go into that: even if facts there are, the view that the truth of a judgment is its correspondence with some fact or other is still untenable. For suppose A judges that Jupiter is

round: call this judgment $J_1$. If A reflects minimally, A will also be able to judge: My judgment that Jupiter is round is true; call this judgment $J_2$. $J_1$ and $J_2$ clearly stand or fall, and indeed both stand, together: they are not made true on two different accounts. Given that $J_2$ is a first-person judgment simultaneous with $J_1$, A who judges $J_1$ needs no further justification, no additional data, to go on to $J_2$. But on the theory of truth as correspondence to facts, $J_1$'s truth would be its correspondence to the roundness of Jupiter, and $J_2$'s truth would be its correspondence to quite a different fact, namely, $J_1$'s correspondence to the roundness of Jupiter. This is good enough reason to reject the theory; all the same, unlike the first theory, it has taught us something: an adequate theory of truth must pass the test that this theory failed, namely that $J_1$ and $J_2$ are made true in the same way and not on different accounts.

I turn then to a much better effort: the so-called redundancy theory of truth, put forward by Ramsey, Ayer, and Arthur Prior. By this theory e.g. 'A truly believes that the Earth is round' would come out as 'A believes that the Earth is round, and the Earth is round': and 'What P.C.49 will tell the Court on oath is true' would come out as 'For any $p$, if P.C.49 is going to tell the Court on oath that $p$, then $p$'. The redundancy theory, as is easily seen, escapes the difficulty that was fatal to the correspondence theory. For 'My judgment that Jupiter is round is true' now comes out as a mere conjunction:

I believe that Jupiter is round, and Jupiter is round. No relation of correspondence comes into the structure of this. If the conjunction expresses something A judges, then so does the second conjunct; the first conjunct, we may say, *professes* the judgment *stated* in the second conjunct; we have no puzzle about two true judgments answering to two facts.

All the same, the redundancy theory too apparently breaks down. I shall not discuss the intricate questions that could be raised over the quantification 'for any $p$' that comes in my second illustration of the theory; whether this quantification is 'substitutional' or 'objectual', and what this difference amounts to. Such identification of variables as we have here is doubtfully legitimate if there would have to be a shift of meaning in any substitution instance of the formula in which the propositional variable is repeated. Frege, as is well known, powerfully argued that there is indeed a shift of meaning (*Bedeutung*) in such cases. If Frege is right, the redundancy theory too must be rejected, or at least is hard to defend.

Why would the identification of variables across such a shift of meaning not be clearly legitimate? An example in which nobody would sensibly dispute the shift of meaning may serve to show this. 'Giorgione was called Giorgione because of Giorgione's size' and 'LittleJohn [the companion of Robin Hood] was called Little John because of Little

John's size' are clearly not to be conceived as got by substitution in '$x$ was called $x$ because of $x$'s size'; it would be absurd to use this last formula for specifying a class {$x$: $x$ was called $x$ because of $x$'s size} to which Giorgione and Little John alike belong. Even without aid of quotes, it is obvious that in these examples the first and the second occurrences of an equiform name have different meanings: what is meant the first time is a man, what is meant the second time is a nickname of his; and this is what makes the formula '$x$ was called $x$ because of $x$'s size' manifestly illegitimate.

Now Frege argues that there is likewise a shift of meaning when e.g. a proper name occurs now in straightforward discourse and now in an oblique context like 'P.C.49 is going to tell the Court on oath that…' – of course not the same shift as we have just considered. The proposition that would occur twice over in a substitution instance of 'If P.C.49 is going to tell the Court on oath that $p$, then $p$' may well contain a proper name; but if so then there is a shift of meaning between the two occurrences of the proper name, and so we have after all only an *apparent* repetition of the proposition it occurs in. Frege's theory of indirect discourse would therefore make the redundancy theory of truth illegitimate in the most important cases: cases like 'What P.C.49 is going to tell the Court on oath is true.'

But *is* Frege right? I do not wish to say he is entirely right; there is something wrong, to my mind, about his positive account of what things proper names come to mean in indirect discourse; but his negative thesis, that there the names do not simply name the objects that they ordinarily name, seems to me solidly established. Let us take an example in every sense more down to earth than his example of the Evening Star and the Morning Star. The Derby winner Running Rein turned out to be a horse Maccabeus disqualified by reason of age for running. In a context like 'Lord George Bentinck discovered that—was four years old' the truth-value might alter according as we inserted the name 'Running Rein' or the name 'Maccabeus': the age of Maccabeus was already well known to the racing fraternity, the age of Running Rein was not. We cannot then read this as a context 'F( )' into whose empty place a name for a horse taken straightforwardly is inserted; for then we could not get 'F (Running Rein)' false while 'F (Maccabeus)' would be true. Whatever else we say, we must say that 'Maccabeus' and 'Running Rein' have not here their straightforward meaning. The meaning they would have, as each naming a certain horse and both of them the same horse, in a context like '— was four years old when he won the Derby'.

I think these considerations rob the redundancy theory of truth of its intuitive simplicity and persuasiveness. Possibly a context in which an inserted proper name would occur straightforwardly at one place and obliquely at another could still be regarded as determining a definite sense for the resulting sentence, a function of the sense of the name: this

was the conception of 'non-Shakespearean' predicables that I suggested in *Reference and Generality*. But the logic of such predicables has not been thoroughly worked out; and the ostensible dissolution of the problem about 'is true' that the redundancy theory offers is paid for at a dear rate if nasty problems of intentionality are left on our hands. The advocates of the redundancy theory, such as Ramsey and Prior, seem to me not to take seriously enough the Fregean case for a shift of a name's use in intentional contexts, so that what it means is no longer the same object as when it is used straightforwardly.

It is worth remarking that Frege himself held what might be called a partial redundancy theory, for 'is true' in certain contexts. He held that the thought expressed by a sentence S was the same as the thought expressed by the longer sentence [The thought that S is true] , where [the thought that S] is of course for him a complex designation of a thought. On this view 'is true' would be *predicable* of thoughts. As we may see at the beginning of the essay 'Thoughts' (*Der Gedanke*), Frege had some discomfort about this account of 'true' in his old age, though he held on to this redundancy theory *faute de mieux*.[i]

Even this limited redundancy theory is open to doubt; in fact Frege's other doctrines, combined with this one, generate an actual antinomy. Frege holds that if a proper name N answers to an unsuccessful identification, then a sentence [F(N)] in which N is meant to be straightforwardly used is neither true nor false, though it still has sense. I accept this doctrine and am ready to defend it. (One class of case that to my mind speaks for it occurs when N is misconferred on two individuals of a kind, e.g. two men, who have been muddled together by the speaker; in this case the speaker literally does not know what, i.e. which one, he is talking about, and his statements are, as lawyers say, void for uncertainty.) But it still holds, even if N fails to name, that the sentence [F(N)] can get over an impression of how things are and others may come to share this impression. Frege would then say that [the thought that F(N)] designates a definite thought; N's vacuousness does not prevent this, because in such oblique contexts N has not its straightforward use e.g. to designate a man. (The thought that Santa Claus descends chimneys is shared by many children, and to recognize which thought it is we need not take 'Santa Claus' to name a man.) But then if 'is true' is predicable of thoughts, the sentences [F(N)] and [The thought that F(N) is true] will *not* express the same thought; the first will express a *truth-valueless* thought when N fails to be a name, the second will say of this truth-valueless thought that it is true, and will thus express a *false* thought. Frege's various insights thus turn out to be partly deceptive.

Dummett resolves this antinomy by proposing to reject the view that in the contexts 'It is true that…' or 'It is false that…' the embedded sentence has its indirect meaning: the sentence would on the contrary

stand for a truth-value, just as if the sentence or its negation were free-standing. This proposal would need modification for sentences with empty proper names in them; but we could now say that just as 'Santa Claus is lazy' has sense but no truth-value, so also the results of embedding this in 'It is true that – ' and 'It is false that – ' have no truth-value; for then 'Santa Claus' does not by such embedding get carried over to its indirect meaning, and in non-oblique use it is an empty name. But we need not consider these empty-name cases to find Dummett's proposal unacceptable. As grammatical subject of 'is true' or 'is false', 'that for no non-zero numbers $x$, $y$, $z$, $n$ do we have $x^{n+2} + y^{n+2} = z^{n+2}$, is replaceable *salva veritate* by the nickname 'Fermat's last theorem'; just as this replacement can be made in other contexts, where Dummett would take both expressions to stand for thoughts. But on Dummett's proposal there would be no indirect meaning of words in the *that* clause when it stood as grammatical subject of 'is true' or 'is false'; we have rather a clause standing for a truth-value. So what is meant by the nickname 'Fermat's last theorem' when it is standing before 'is true' or 'is false' will likewise be a truth-value! Dummett says 'this is a consequence which it is possible to swallow if one is resolute'. (*Frege: Philosophy of Language*, p. 382). When the Duke of Wellington, no coward assuredly, had ingested an over-hot potato, he did not show the resolution Dummett commends: he promptly spat it out on his plate, and remarked to his hostess the unwisdom of swallowing it.

To my mind, Dummett and Frege both go wrong about the semantic role of sentences. Both hold that a sentence stands to *something* in the same kind of relation as a name does to what it names: not always to the same thing – to a truth-value in straightforward use, to a thought (*Gedanke*) in oblique contexts. The term 'truth-value' may have caused misunderstanding; it does not here mean the circumstance of a given sentence S's being true or, as the case may be, false (who could doubt that 'there are' the two truth-values *true* and *false* in this sense?); rather it is held that all true sentences straightforwardly taken stand in a namelike relation to one entity, the True, and all false sentences, to another entity, the False.

Dummett modifies Frege's theory just to the extent of saying that the True and the False are *not objects*. This modification is quite ineffectual. It is surprising that he should think it effectual: for in criticism of what I once wrote he justly objected that whether numerals are names cannot be made to depend on whether numbers are objects; if numerals logically behave like non-vacuous names, then they *are* such, and then numbers must be recognized as what numerals name. Fine; but equally the primary question is whether sentences play a kind of naming role; if so, then it is futile to try to discriminate this from the role of names properly so called by saying that what are here named are not *objects*.

What we should say is that sentences simply are not names, do not stand to *anything* in a namelike relation; neither in their straightforward use (freestanding, or as truth-functional components) nor in their oblique use. We have to take into account the *logic* of duality. To grasp this difficult notion intuitively, we need to recognize that (if we consider just statement-making sentences) a language alternative to ours is possible in which a sentence equiform to an English sentence says the contradictory opposite of what the English sentence says. Let us call the alternative language in which this is done 'Unglish'. The Unglish sentences 'It is not raining' and 'It is raining' will then respectively render the English sentences 'It is raining' and 'It is not raining'; so 'not', as in English, can serve to form a contradictory for a sentence; the Unglish for 'not' is 'not'.

This way of explaining duality originates with *Tractatus* 4.062-.0621. Max Black in his commentary misunderstands the matter: he in effect supposes that in English 'It is raining', Unglish 'It is not raining', we have Unglish 'not' translating the *absence* of 'not' in English! This is a peculiarly piquant example of what I have called the *cancelling-out fallacy*: the error of supposing that if equiform expressions are cut out from two sentences which *as wholes* have the same sense, then what is left must have the same sense; here, what is left on one side is what linguists call the zero morpheme, i.e. *nothing*. Naturally if this first step were right, the idea of a dual language Unglish would dissolve into incoherence; but the step is wrong, and as I said, the Unglish for 'not' is 'not'.

In working out which pair of expressions are mutually dual, we are so to say constructing an English-Unglish (or Unglish-English) dictionary; we want to find a set of pairs of expressions such that if we replace each expression in the set by its mate we get a negation of the original sentence. The theory of duality has been extensively worked out. Any propositional part within a sentence is dual to its negation, as the whole sentence is. The connective 'either... or...' is dual not to 'neither... nor...' but to 'both... and...'; 'some' and 'every' are dual to one another. Names are self-dual, for in the contradictory we shall still be mentioning the same things by saying the opposite about them; and by the simplest key of translation, common nouns like 'horse' in quasi-subject positions (e.g. in a phrase 'every horse' or 'some horse') will also be self-dual; this speaks in favour of the old doctrine that in such places these words too are names. Predicables, on the other hand, are dual to their negations. Take the sentence 'Every horse galloped'. The quantifier 'every' is dual to 'some'; 'horse' is self-dual; 'did not gallop' is dual to 'galloped'; so putting the bits together we get 'Some horse did not gallop', contradictory, as it ought to be, to our initial sentence.

For our purpose, two special cases of duality are important. First, what is dual to a definite description or other complex singular

designation? Frege treated such expressions as complex names; by that reckoning they would be self-dual. I do not believe there are any complex names; a name needs no internal structure in order to be a name, so any structure it happens to have physically is irrelevant to its sense. But we *nearly* get self-duality for complex singular designations. We may distinguish two workable definitions of [*F* (the one and only *A* that is *G*)], which may be given the following semi-English explanations:

       (i) [Just one *A* is *G*, and *F* (that *A*)]
       (ii) [If just one *A* is *G*, then *F* (that *A*)]

     Let us write [*F*(the *A* that is *G*)] for the first, and [*F*(the *A* that is *G*)] for the second. Then if [*F*()] and [*F'*()] are contradictory (and thus mutually dual) predicables, which yield contradictory propositions when one name is inserted in their argument-places, [*F*(the *A* that is *G*)] and [*F'*(*the A* that is *G*)] work out as dual to each other; and the difference between the two mutually dual readings of [the one and only *A* that is *G*] becomes unimportant when the truth of [Just one *A* is *G*] is guaranteed.

     The other important case of duality is the dual to a proposition-forming functor with a proposition as argument. Let $\varphi$ be such a functor: the dual to $\varphi$ is [¬$\varphi$¬] – the successive application of negation, $\varphi$, and negation again. At any rate, this will be so if we assume, like Frege, that double negation of a proposition does not alter its sense. For by our rule [(¬$\varphi$ ¬)(¬P)] will be dual to [$\varphi$ P]; now [(¬$\varphi$ ¬)(¬P)] is the same as [(¬$\varphi$)(¬¬P)], which has the same sense as [(¬$\varphi$)P]; and this last is just another way of writing [¬($\varphi$ P)], which is dual to [$\varphi$ P] as it should be.

     The law of double negation is already disputed in some quarters, let alone the Fregean principle that double negation does not change sense. Obviously I cannot here argue the question. Following Elizabeth Anscombe, I want to say that two propositions' being one another's negations is like two correlative terms' being one another's converses; that double negation no more alters the sense than 'Cnv' iterated does; that the idea of inherently negative propositions, whose contradictories are inherently positive, is as empty as the idea of a class of inherently converse relative terms, which are converses of (shall we say) basic relative terms. I think too that what Intuitionists are after could be better secured by restricting the use of the dilemma pattern of argument (the *vel*-elimination rule) rather than by rejecting the laws of double negation and excluded middle. It must for now suffice to have said this; henceforth I take for granted Frege's principle that double negation does not alter the
sense.

Let us now consider how duality works with oblique occurrences of propositions. There are two rival theories: one conformable to Frege's or Dummett's views, the other to Arthur Prior's. Take the following sentence:

Jones is informed that Smith has been in prison.

By the Frege-Dummett account, 'that Smith has been in prison' is a complex designation of a certain thought; since thoughts, for Frege, are individual pieces of information (two sentences conveying the same bit of information convey the same thought), we may use the paraphrase:

Jones is apprised of the piece of information that Smith has been in prison.

Now let us consider how duality will work. The string of words following 'of', as in the simpler case of ordinary definite descriptions, will admit of two mutually dual readings, which as before we may distinguish by using Roman and italic type for the definite article. The name 'Jones' will be self-dual; 'is apprised of' will be dual to 'is not apprised of'; so, piece by piece, the dual of the whole sentence works out as:

Jones/is not apprised of/*the* piece of information that Smith has been in prison.

Since for argument's sake we may suppose there is no question of our words' relating to *no* piece of information, or to more than one, we need not bother about the difference between the 'the' designation and the '*the*' designation; the duality thus far seems to work out satisfactorily – but it only seems to. For let us now consider the duals of the expressions within the *that* clause. The predicable 'has been in prison' is dual to its negation 'has not been in prison'; and even if we agree with Frege (as I say we *should*) that 'Smith' here has not a straightforward meaning, standing for a man, that is no reason for not treating it as still a self-dual name. (Frege of course would take it to name the relevant ordinary sense of the name 'Smith'.) If we take *these* dualities into account, we get:

Jones/is not apprised of/*the* piece of information that/ Smith/has *not* been in prison.

Something has gone badly wrong. Of course some further complications of the theory could be devised to avoid a crack, but they would to my mind be just adhocus-pocus.

Let us now look at Prior's rival analysis. Prior treats 'Smith has been in prison' as still being a proposition even in oblique constructions, and correspondingly treats 'Jones is informed that – ' as a proposition-

forming functor with a single propositional argument. By our general rule, the dual of the functor will be the product of negation, this functor, and negation again:

> It is not the case that/Jones is informed that/it is not the case that

Now let us turn to 'Smith has been in prison'. The dual of this, as before, will be 'Smith has *not* been in prison'. Putting the pieces together, we get:

> It is not the case that/Jones is informed that/it is not the case that/Smith has *not* been in prison.

But this, by the Fregean principle that double negation makes no difference to the sense, will reduce to:

> It is not the case that/Jones is informed that/Smith has been in prison,

which is, as it ought to be, the contradictory of the proposition we first thought of. This result comes out without any new apparatus or adhocus-pocus: as Frege might have said, *bei der richtigen Auffassung kommt Alles in Ordnung.*

I believe, then, that a theory on Prior's lines could be coherently worked out and would be manifestly superior to one on Frege-Dummett lines. I do not think Prior's own theories, historically speaking, were quite satisfactory. He did not accept the need for distinguishing as Frege did the straightforward and the oblique meaning of names; and this pushed him step by step, as Russell had been pushed, towards a very restrictive view of real proper names, proper names for which Frege's problem would not arise. But as I said before, allowing that the name 'Smith' in an oblique context does not straightforwardly mean the man Smith, we are not then estopped from still treating it as self-dual. I must leave to others the exploration of the possible escape-routes whereby the dual of 'Smith' in oblique contexts would be some other expression, or 'Smith' would have two different oblique uses to make dualities work out right.

My view is, then, that sentences are not names and are nothing like names. Whether a sentence occurs straightforwardly or obliquely within another, it neither names nor has any namelike relation to anything whatsoever, whether a truth-value or a thought or a state of affairs or what you will. Names admit of no duality of significance: a name either names something, or simply fails to name and thereby becomes semantically futile. Sentences are essentially dual in significance: what a true sentence points towards is what its contradictory points away from; there are two opposite semantic relations involved. A false sentence is

not like an empty name; for it, unlike an empty name meant to be taken straightforwardly, can be an integral part of a sentence with truth-value. 'Pointing towards' and 'pointing away from' are of course metaphors, but what they are metaphors for cannot be informatively explained; an inchoate understanding of these relations is involved in all informative discourse, and this understanding can be clarified, or sharpened, by logical and philosophical training, but there can be no question of analysis or explicit definition.

Here I had better quickly clear up the puzzle we had about such apparent names as 'Fermat's Last Theorem'. What we get here, I hold, are not names but mere abbreviations for *that* clauses: the apparent name 'Fermat's Last Theorem' is not related to the clause 'that for no non-zero integers $x, y, z, n$ do we get $x^{n+2} + y^{n+2} = z^{n+2}$' as a proper name is to a definite description. A proper name is not an abbreviation for a definite description (I have argued this elsewhere, and shall not do so here); but expressions like 'Fermat's Last Theorem' are *mere* abbreviations. What a sentence signifies (or better: *how* a sentence signifies things to be) can be signified *only* by a complex sign; thus medievals rightly spoke here of *complexe significabilia*; an abbreviation is a mere proxy for such a complex sign. Were it otherwise, 'I believe Pop and he believes Pip' could be a plainer way of conveying the different belief-relations in which I and he stand; it would be a *pis aller* to describe the entity Pop by the rigmarole 'that the Earth is round' and the entity Pip by the rigmarole 'that the Earth is flat'. (A reduction, I hope, of latent nonsense to patent nonsense.) And so we may forget about Dummett's puzzle which thing is meant by 'Fermat's Last Theorem' in 'Fermat's Last Theorem is true' and his heroic solution that here it stands for a truth-value, though elsewhere for a thought.

To what, then, do a pair of contradictories alike relate – one pointing towards it, and the other away from it? There are many reasons for rejecting the answer: A fact. Facts, as I said, are *entia non grata* because of their uncertain individuation; and moreover I think the appearance of the construction 'the fact that...' always points to a need to split up the sentence, so as to exhibit the content of the *that* clause in a separate assertion (see my Logic Matters, pp. 21-23 and 259-261). But moreover it is easy to wield Ockham's razor drastically here. Frege, with only one semantic relation for sentences used straightforwardly to bear, could cut down what sentences relate with to just two objects, the True and the False. If we recognize a duality of semantic relations for sentences, we can account for their semantics in terms of just the True: all true propositions point towards this, all false ones away from it.

Here, as I said, no informative analysis is possible: similes may help. In a world where all roads lead to one place, to Rome, let us say, travellers who face Rome will all meet there if they follow the road, travellers whose backs are towards Rome will be scattered to all

quarters. The different roads to Rome and from Rome correspond to the different senses of true propositions or again of false propositions; contradictory opposition is represented by travellers going opposite ways along the same road. On a round Earth indeed travellers all going straight away from Rome would meet at Rome's antipodes; but it is not a logical truth that the Earth is not flat, and there are no logical antipodes.

Another simile brings out a little more. A parallel beam is turned by a convex lens into a convergent pencil of rays passing through a real focus; a concave lens would turn the beam into a divergent pencil of rays which when projected backwards meet in a virtual focus. At the real focus there is an actual concentration of energy, physically related to the beams that meet there; there is nothing at the virtual focus physically relevant to the divergent rays. The difference between the relations of the convergent and the divergent pencil to the focus here represents the difference between true and false propositions; the difference between two rays in the same convergent or divergent pencil represents the difference in sense between two true or two false propositions. And as we shall see, I wish to say that the True is *causally* related to true saying and thinking, as it is not to false saying and thinking: in the model this is represented by the different *physical* relation of a real and a virtual focus to a pencil of rays *geometrically* passing through the focus. Obviously any such simile must limp somewhere; it would be idle to seek for analogues of the originally parallel beam or of the two lenses.

At this point someone may be reminded of the doctrine taught by Augustine and Anselm, that all true saying and thinking relates to one Truth, which is God. This is no accident; I do myself believe that the True, the goal and focus of all true saying and thinking, is indeed the living, the only God. But I shall not here try to prove this; I shall indicate what steps would have to be climbed and how arduous the climb would be.

Augustine had an all too easy argument for the eternity of truth: if we tried to deny this, we should get 'It was once true, or will eventually be true, that there is no truth', which is absurd. It is providential that Augustine did not devise this argument in his Manichaean days, or he might have found his logic forcing him to remain a Manichee, as follows: 'Falsehood must be eternal, for otherwise we should get:

It was once false, or it will be false, that falsehood exists

and this is absurd'. Something we might indeed characterize as logical Manichaeism is Frege's position: his setting up of the True and the False as two objects of reference implicitly recognized in all thinking, even that of the sceptic.

We are delivered from logical Manichaeism by the recognition of duality. But to show that the True is a living God we must show that the True can be credited with understanding and will; and moreover *causes* true thinking and saying in this world, besides *being positively signified* by them. This would be a long story.

As regards understanding, my first step (and I believe this is possible on somewhat similar lines to those in my *Three Philosophers*) would be to expound and defend Aquinas's doctrine of *esse naturale* and *esse intentionale*: the doctrine that a set-up in the world and its mental representation differ not by a descriptively capturable difference between two existent things, but by the different *manner* of existence ('existence' here being used in a sense *not* expressible by the quantifier 'for some *x*'). This could be used, I think, to explain our earlier result about the relation between a man's judging that Jupiter is round and his judging that he himself so judges. By Aquinas's account (*Summa Theologica* Iaq.16 art.2) A in judging that Jupiter is round brings into actuality in himself an intentional instance of the same form *round* as exists *naturally* in Jupiter if A judges truly; in bringing it into actuality, A is aware of so doing, and *eo ipso* judges its agreement in form (*conformitas*) with the roundness existing in Jupiter, and this constitutes a simultaneous judgment that his judgment about Jupiter is true. Aquinas considers only very simple cases, and we need to supplement his account by some recursive procedure to cover the truth-conditions of more complicated judgments; but I think this is at least the first step towards a correct account.

From this I would go on to Aquinas's notion of *entia actu intelligibilia*, entities whose very life is a thought by each one of himself; such are not we, although we too do think, but such, Aquinas holds, are the immortal spirits, good and evil; and such, I would argue, is the True. If understanding can be ascribed to the True, so can will; for 'will' does not mean a faculty of eliciting mental states called volitions, but a mode of *causality*, voluntary causality, that is proper to beings who have discourse of reason. And it is by that will, I should argue, that the True *brings about* all true thinking and saying, like an artist creating multitudinous self-portraits. This would be terrible egotism on the part of a human artist; but then there is nothing better than the True to be represented, and no representations can exhaust its riches.

All this is only a sketch of an argument I have partly expounded elsewhere. But I think I have shown in the present paper that there is solid reason to believe that all our true saying and thinking points to one object, the True. To do just that is all we are here for. Christ said that he was born and came into the world to bear witness to the Truth; unless in our small measure we too do that, we are worthless; our life has failed like a seed that never germinates. In comparison with this goal, how paltry it seems to devote oneself to the godling of some modern thinkers:

a godling changeable, and ignorant, and liable just as we are to passions like anger and grief and access of joy! I have not proved that the True is God, but I will worship nothing else: if the True is not God, there is no God.

---

[i] p. 6 of the Blackwell translation of *Logische Untersuchungen*.

*Truth and God*

GRAHAM OPPY

# TRUTH AND GOD

## GRAHAM OPPY

GEACH (1982) is distinctly odd. On the one hand, it contains a fairly breezy dismissal of some serious theories of truth; on the other hand, it recommends some positive views about truth that clearly deserve much shorter shrift than the summarily dismissed theories.

On the first hand, Geach examines, and finds fault with, 'the speech-act theory of truth', 'the correspondence theory of truth', and 'the redundancy theory of truth'. On the second hand, Geach asserts—and hints at a defence of—the following four theses:

1. All our true saying and thinking points to the True (a.k.a. God).
2. The life of the True is its thinking of—i.e. about—itself.
3. The True has voluntary causality because this is proper to beings that have discourse of reason.
4. By its will, the True brings about all true thinking and saying.

Geach begins by locating the need for a theory of truth in locutions like 'Whatever the Pope says is true'. According to Geach, 'we cannot extract something to the same effect … no longer involving mention of truth' (84). On its face, that seems wrong; consider: 'For any p, if the Pope says that p, then p'. Various authors, from Quine (1970) to Horwich (1998), have observed that the truth predicate serves a logical need that might be filled in other ways: e.g. propositional quantification (as above), or infinite conjunction ('If the Pope says that grass is green, then grass is green; and if the Pope says that grass is red, then grass is red; and if the Pope says that snow is cold, then snow is cold; …'), or generalisation over a sentential schema ('we should accept all instances of the schema: *if the Pope says that p, then p*'). Nonetheless, it is clearly correct that any adequate theory of truth must accommodate locutions such as the one that Geach mentions.

Speech-act theories of truth are tailor-made for cases like 'Everything the Pope says is true'. As Geach observes, on speech-act theories of truth, this sentence would be assimilated to a sentence containing an appropriate performative verb, e.g. 'I corroborate everything that the Pope says'. However, as Geach also observes, there are other sentences that appear to resist this kind of analysis. Consider, for example: 'What Assange said is a scandalous revelation if true'. (The kind of objection

that Geach is deploying here has many uses—see, for example, Geach (1960) (1965).) I agree with Geach that this objection is crippling for speech-act theories of truth.

Geach is also very short with correspondence theories of truth. According to Geach, the following objection is decisive: "Suppose A judges that Jupiter is round: call this judgment $J_1$. If A reflects minimally, A will also be able to judge: My judgment that Jupiter is round is true; call this judgment $J_2$. $J_1$ and $J_2$ clearly stand or fall, and indeed both stand, together: they are not made true on two different accounts. … But on the theory of truth as correspondence to facts, $J_1$'s truth would be its correspondence to the roundness of Jupiter, and $J_2$'s truth would be its correspondence to quite a different fact, namely, $J_1$'s correspondence to the roundness of Jupiter. This is good enough reason to reject the theory." (84) It is very unclear what Geach is arguing here. On the one hand, it seems evident that 'Jupiter is round' and 'I judge that Jupiter is round' are made true on two different accounts (just consider how matters stood in 1934, when my father was but five years old). On the other hand, it seems no less evident that 'I judge that Jupiter is round' and 'I judge that my judgment that Jupiter is round is true' are made true on two different accounts (since the former concerns my judgment of Jupiter's shape, and the latter concerns my judgment of the correctness of my judgment of Jupiter's shape, and, at the very least, these two kinds of judgments could come apart with the passage of time—I could come to judge that my previous judgment of Jupiter's shape was incorrect). Note that this criticism can accept Geach's contention that his two judgments—whatever exactly their content might be—'stand or fall together': at a given point in time, I will judge that my judgment that Jupiter is round is true just in case I judge that Jupiter is round, but this is insufficient to establish that there are not two distinct judgments being made. Since Geach says nothing else against correspondence theories, I do not think that we should hastily suppose that they can henceforth be ignored. (See David (1994) and Horwich (1998) for further discussion of the prospects for correspondence theories of truth.)

Geach gives far more attention to 'the redundancy theory of truth', which Geach claims would render our key sentence as 'For any p, if the Pope says that p, then p'. Against this view, Geach claims that there is enough that is right in Frege's theory of indirect discourse to make this rendition of our key sentence 'illegitimate' (85). The difficulty that Geach discerns is that, if Frege's theory of indirect discourse is correct, then, for many substitution instances of 'if the Pope says that p, then p', there will be a 'shift in meaning' in what is substituted for the first and second occurrences of 'p'. While there is quite a bit to say about this, I

think that is suffices to observe that the argument that Geach takes to establish that there is enough right in Frege's theory to justify this conclusion is far from compelling. On the one hand, there are many contemporary theorists—e.g. Salmon (1986) and Soames (2005)—who reject the claim that there is a 'shift in meaning' in what is substituted for the first and second occurrences of 'p'. On the other hand, there are more sophisticated understandings of variables and quantifiers that may accommodate the alleged 'shift in meaning' (see, for example, Oppy (1992)). So Geach's objection to what he calls 'the redundancy theory of truth' is plainly less than airtight.

Geach gives even more attention to a 'partial redundancy theory of truth' that he claims to find in Frege: namely, the view that the thought expressed by a sentence S is the very same thought that is expressed by the sentence 'The thought that S is true'. Geach notes that, given Frege's view, that sentences containing empty names in extensional contexts lack truth values but not senses, there will be cases in which a sentence S lacks a truth value even though the sentence 'The thought that S is true' does not—whence, by Frege's own lights, the 'partial redundancy theory of truth' cannot be correct. While this criticism of Frege seems correct, we have already noted that one might well elect to go in for more wholesale rejection of the Fregean package than Geach himself countenances (both here and in his subsequent discussion of Dummett's attempts to repair this particular puncture).

Having satisfied himself that no extant theories of truth are satisfactory, Geach proceeds to some ground-clearing in preparation for the statement of his own account of truth. The main claims for which he argues are: (1) sentences are not names; (2) sentences do not stand to anything in a name-like relation; (3) sentences are essentially dual in significance; (4) what a true sentence points towards is what its contradictory points away from; (5) there is a duality of semantic relations for sentences; (6) all true propositions point towards the True; all false propositions point away from the True.

There are many questions that one might ask about this 'ground-clearing'. In particular, one might wonder what could motivate invocation of 'the True', once we have given up the idea that there are things to which sentences stand in name-like relations. For the purposes of logic and systematic semantics, we have good reason to invoke 'truth-values'. If—following Frege—we insist on reifying or objectifying 'truth-values', then we end up, at least, with a commitment to 'the True' and 'the False'. (We here—and henceforth—set aside worries about whether we need more 'truth-values' in order to cope with sentences that are

neither true nor false, etc.) If—apparently following Geach—we say that it is a mistake to reify or objectify 'truth-values', then it seems that we have equally good reasons to dispense with both 'the False' and 'the True': we can still say, in the absence of both, that true sentences 'have' the truth-value *true* and false sentences 'have' the truth-value *false*.

Furthermore, even if we accept Geach's contention that we do have good reason to accept the existence of 'the True', it is not clear that we would not then have equally good reason to accept the existence of 'the False'. Geach says that all true propositions point towards the True and all false propositions point away from the True. But it is only by pointing towards something else that something can point away from a given thing. True enough, you might ask me to point away from you: but, if I do not point at something else, then I am merely pretending to point away from you. In order for it to be the case that false propositions point away from the True, there must be other things at which those false propositions point—and, in the nature of the case, the most natural hypothesis would surely be that they point towards the False. (True enough, Geach says that 'pointing towards' and 'pointing away from' are metaphors that cannot be informatively explained or informatively analysed. But I am not here asking for explanation or analysis; rather, I am just following Geach's metaphors where it seems to me that they most plausibly lead.)

Finally, even if one accepts that pairs of contradictory sentences are alike related to single entities, one might well prefer an answer that Geach rejects: namely that pairs of contradictory sentences are related to facts (or true propositions, or the like). Geach dismisses facts on the grounds that (a) their individuation is uncertain; (b) the construction 'the fact that …' has misleading surface features that disappear under proper analysis; and (c) postulation of facts to play the role assigned to them multiplies entities way beyond necessity (since we could clearly get by with just the True and the False, if not with just the True).

While it must be conceded that some theorists still harbour doubts about facts—see, for example, Neale (2001)—it is hardly in doubt that one can provide clear individuation criteria for facts, and that one need not accept Geach's construal of 'the fact that …'. Moreover, one might naturally think that the way in which sentences get to point towards the True or the False is precisely by way of their correlation with facts, or true propositions, or truths, or the like. Because he accepts so much of the Fregean machinery, Geach happily accepts that there are thoughts – structured senses – that are correlated with sentences: but one could surely raise the same kinds of worries—about individuation criteria for

thoughts and construal of 'the thought that ...'—that Geach raises in connection with facts; and one might surely think that 'representations' of thoughts could fairly readily be reinterpreted as 'representations' of propositions ('representations' of true thoughts could fairly readily be reinterpreted as 'representations' of facts).

While there are further objections that might be raised against Geach's 'ground-clearing', it seems to me that there is ample reason to think that we have not been given good reasons to accept his claim that all true propositions point towards, and all false propositions point away from, the True.

What about the further theory that Geach erects on the basis of his 'ground-clearing' propositions; in particular, what about his identification of the True with God. I confess to finding this barking mad. Geach says: all our true saying and thinking points to God. Against this, it seems to me to be evident that Theists and Naturalists ought to be able to agree on a theory of truth. Here, the theories that Geach dismisses point the way: there is no obvious reason why Theists and Naturalists could not agree on the speech-act theory, or the correspondence theory, or the redundancy theory (or the coherence theory, or the identity theory, or any of the other serious theories of truth that have done the rounds in the past fifty years). And, of course, Naturalists think that it is true that there is no God. But how, exactly, could our saying truly that there is no God point to God?

REFERENCES

David, M. (1994) *Correspondence and Disquotation: An Essay on the Nature of Truth* Oxford: Oxford University Press

Geach, P. (1982) 'Truth and God' *Proceedings of the Aristotelian Society*, Supplementary Volume 56, 83-97

Horwich, P. (1998) *Truth*, second edition Oxford: Clarendon

Neale, S. (2001) *Facing Facts* Oxford University Press

Oppy, G. (1992) 'Why Semantic Innocence?' *Australasian Journal of Philosophy* 70, 4, 445-54

Quine, W. (1970) *Philosophy of Logic* Englewood Cliffs: Prentice-Hall

Salmon, N. (1986) *Frege's Puzzle* Cambridge: MIT Press

Soames, S. (2005) *Reference and Description* Princeton: Princeton University Press

Wright, C. (1992) *Truth and Objectivity* Cambridge: Harvard University Press

JANE HEAL & JULIAN DODD

## BIOGRAPHY

Jane Heal studied for her first degree in Cambridge, reading History for two years and then Philosophy (or 'Moral Sciences' as it was then called) for another two years. She also took her PhD at Cambridge, working on problems in the philosophy of language. After two years of post-doctoral study in the US, at Princeton and Berkeley, she was appointed to a Lectureship at the University of Newcastle on Tyne. Having taught there for ten years, she moved back to Cambridge, where she is now Emeritus Professor and a Fellow of St John's College. She is the author of *Fact and Meaning, Quine and Wittgenstein on Philosophy of Language* (Blackwell 1989) and *Mind, Reason and Imagination* (CUP 2003). She was elected a Fellow of the British Academy in 1997 and president of the Aristotelian Society in 2001.

"The Disinterested Search for Truth" was originally published in the *Proceedings of the Aristotelian Society, Volume LXXXVIII* (1988).

Julian Dodd is a professor of philosophy at the University of Manchester. He works in a number of areas of philosophy, notably the philosophy of language, metaphysics, the ontology of art, and the philosophy of music. Preferring to follow his nose, rather than follow a ruthlessly focussed research programme, he has chosen to write on a variety of topics that have struck him as interesting and important: topics including the nature of truth, the ontological status of works of music, and what it is to perform musical works authentically. Significant publications include the monographs, *An Identity Theory of Truth* (2001) and *Works of Music: An Essay in Ontology* (2007): books produced even though he probably spent more time on his bike than was good for him.

# The Disinterested Search for Truth

## JANE HEAL

# THE DISINTERESTED SEARCH FOR TRUTH

## JANE HEAL

Truth is generally thought to be a Good Thing. We aim (or should aim) at speaking the truth and acquiring true beliefs and should eschew uttering falsehoods or acquiring false beliefs. How are we to understand these seeming truisms? I want to suggest that they are liable to misconstrual. We operate with a rough and ready distinction between evaluative and non-evaluative terms and tend to hear the above remarks as placing 'true' in the evaluative bag. My claim is that, insofar as we have any grip on this evaluative/non-evaluative distinction, 'true' and 'false' are wrongly so placed. They do not function either with respect to utterances or beliefs as do other cases of 'evaluative' terms – for example 'just', 'wicked', 'enjoyable', 'delicious', 'beautiful', 'hideous', 'interesting'. What all these terms have in common, and what I shall take to be at the centre of their 'evaluative' status, is that they have the role of making actions intelligible by helping to show us what an agent saw in those actions. I shall maintain that 'true' does not fulfil this function. Rather, whenever we find a case where what seems to be involved is the pursuit or promotion of truth a closer look will always show us that the agent's project can be better described in another way which does not involve the mention of truth. Truth is sometimes said to be 'transparent' and we could use that image to put the point: the transparency of truth allows whatever valuable features there are in a situation or project to shine through but does not itself contribute anything of substantive value. Taking the contrary view can, I suggest, support a diverse range of mistakes both in philosophy of language and in our conception of the importance of scientific and other academic enquiries.

For the sake of brevity I shall talk mainly of truth. What I have to say can be applied *mutatis mutandis* to falsehood as well. I shall comment separately on those occasions where the two may seem to diverge.

'True' is used as a predicate of (a) utterances, (b) judgements or beliefs, (c) propositions, regarded as abstract objects which can be the content or object of (a) and (b). In any sense in which we can promote or bring about truth (c) is clearly not at issue. So I shall confine my

discussion to (a) and (b), concentrating on utterances in Part I and on beliefs in Part II.

<div align="center">I</div>

In a number of recent discussions of meaning we find mention of the idea that when a person understands or knows the meaning of some word or sentence he or she is thereby committed to some pattern of use of that word or sentence.[1] Let us take Kripke's familiar example. We are to imagine someone who is asked 'What is 68 plus 57?'. We are to assume that 68 plus 57 is 125, that the person knows what '68', '57' and '125' mean. Taking these things for granted, Kripke asks us to contemplate what is involved in the person's taking 'plus' to mean addition. His claim is that it involves a normative element. The claim that 'plus' means addition for this person is not to be linked to the descriptive claim that the person *will* (very probably) answer '125' to the question but to the claim that he or she *ought* to answer '125'.

Some writers have reservations about putting the point quite so bluntly. For example Blackburn remarks that what a person means at one time does not oblige him to say something at another time if he has changed the meaning of his words in the interim. And McDowell points out that the talk of 'pattern' in this context is metaphorical at best since 'what it is correct to say with the use of a given concept, even supposing a given state of affairs one aims to describe, depends upon what other concepts one chooses to express in the same utterance'.[2] But the idea that some obligation to behaviour is bound up with acceptance of something as having a certain meaning persists through these refinements. The basic thought seems to be that if a noise has a certain meaning then it will be true (or at least a component of something true) under certain circumstances and that truth is the correct, right, proper or obligatory thing to say.

I would like to suggest however that we should not merely be cautious about how exactly we formulate the claim but should reject it

---

[1] S. Kripke, Wittgenstein on Rules and Private Language, Blackwell 1982. J. McDowell, 'Wittgenstein on Following a Rule', *Synthese* 1984, pp. 325-363. S. Blackburn, 'The Individual Strikes Back', *Synthese* 1984, pp. 281-301. C. Wright, *Wittgenstein on the Foundations of Mathematics*, Duckworth 1980.
[2] Blackburn, *op. cit.* p. 287.
McDowell. *op. cit.* p. 359, n. 3.

altogether. In learning the meaning of a word I acquire no commitment to, no reason in favour of, any action whatsoever. I acquire indeed knowledge that such and such an action would constitute a true utterance – and hence that such an utterance would be 'correct' or 'what ought to be said'. But this 'correct' or 'ought to be said' amounts to no more than 'true' in this context. And the truth just as such has no claim to utterance.

This may seem very shocking. But let us consider some cases. Take someone who says something which has, and could be foreseen by him to have, bad results for him – for example he cracks a joke at the expense of someone on an interviewing panel and thus loses himself a job. We may enquire why on earth he made that disastrous remark. If he replies 'Well, it was very funny' (and goes on to convince us that this is indeed so) then we may get some way to seeing the poinlt. The amusing does have some claim to utterance; to indicate that a remark was amusing goes at least some way to making the man's action intelligible, if not defensible. We see what he saw in it and we know now what kind of man he is – one who just cannot resist making a joke. That kind of person is familiar and comprehensible to us.

But suppose that our interviewee's disastrous remark, the one that lost him the job, was not amusing (nor informative on some relevant matter, nor such that the making of it was a matter of religious principle . . .) but merely *true*. For example, he suddenly remarks that his interviewer is grossly fat or informs the panel that the earth is eight and a half light minutes from the sun. Later he tries to explain these *prima facie* bizarre actions by saying that a truth just popped into his head and he could not resist enunciating it. Can we make any sense of this?

That someone was nervous and just blurted something out, that he could not resist the chance of showing off his knowledge – these make sense. But they are cases where we either withdraw the claim that the utterance was in a full sense an action or begin to invoke motivations other than that of merely uttering a truth. One may offer the defensive remark 'Well, it is true' on being accused of hurting someone's feelings with a critical personal comment. But the context here suggests that the assumption is that people ought to know such things about themselves.

The point of these examples is not to show that in many cases other considerations, e.g. those of politeness, relevance or caution, might outweigh our supposed commitment to continuing the pattern of

utterances by producing some true sentence. Anyone would have to admit this. Imagine that we are in hiding from ruthless pursuers who will kill us if they find us; to pass the time we set each other mathematical problems; I may ask you 'What is 57 plus 68?' and you should, of course, say '125'; but just at that instant we hear our enemies approaching our place of concealment. Do you speak? No one is going to suggest that you should. But I am saying something stronger than this, namely that truth by itself has no motivating power at all. What obligations, temptations and justifications there are to speak stem from other features of the situations in which we find ourselves – from the fact that someone needs to know something, that we should show friendliness, that it is our turn to contribute to the discussion and so forth.[3] In Kripke's original example, surely the fact that a question has been posed as well as the fact that 'plus' means addition is necessary to us finding it so natural to say that the hearer *ought* to say '125'.

What about falsehoods? I wish to maintain similarly that there is nothing wrong in uttering the False. This sounds terrible. Am I condoning lying? Not at all. The claim is compatible with the strictest Kantian standards. But Kant would not have condemned the telling of a fairy story beginning 'Once upon a time there was a giant with three heads' merely on the Russellian grounds that this sentence was one of a set of falsehoods. And telling fairy stories is not a case where a small amount of some value is sacrificed but we put up with it for the sake of a greater good, or some *prima facie* wrong is committed.

What is wrong about lying is that it constitutes betrayal, harm, abuse of trust, manipulation or the like. In particular cases the vehicle of the betrayal etc. will be the utterance of a false sentence. And we must acknowledge that the falsity of the sentence in these cases is no mere adventitious feature but is central to the nature of the bad thing done. But the relation here is the same as that between, say, giving a child sweets and distorting her understanding of the relations between virtue, reward and love. I may be doing this bad thing in giving her sweets, and the sweetness of the sweets is essential to the matter if that is what is going on. But for all that sweets are not intrinsically wicked.

---

[3] For a similar claim see D. Davidson, 'Communication and Convention', *Inquiries into Truth and Interpretation*, Clarendon Press 1984, p. 268.

Let us note briefly that these reflections have implications for current discussions of language and meaning. It is not impossible to read into Kripke some such line of thought as the following: 'True' is a normative or evaluative notion in that to say that a sentence is true is to say that it ought to be uttered. Statements about meaning entail (together with other non-evaluative premises) statements about truth. Therefore these statements about meaning are evaluative too. But no mere facts (about images, feelings or formulae in the head) can make it the case that I ought to utter some noise. What then am I doing when I say that something is true? I am not describing a fact, I am expressing my approval or feeling of satisfaction (naturalness, familiarity) with that utterance. And when I ascribe a certain meaning to someone else's utterances I am expressing my approval of him as a member of my linguistic community, showing that I regard him as worthy of being responded to with feelings of satisfaction or otherwise.[4]

On this interpretation of Kripke he is seeing a likeness between Hume and Wittgenstein not only with respect to the structure of their respective views about causation and meaning but even more importantly with respect to their views about fact and value. The logical interest and distinctiveness of the concept of meaning boils down to no more than its being an evaluative and so not a factual matter. Whatever we may think of this as a reading of Kripke (let alone Wittgenstein) it is clear that some have indeed taken the Kripkean thesis this way.[5]

But if the earlier claims of this paper are right, the starting point for the whole line of thought is mistaken. Ascription of truth is not bound up with provision of motivation for action. Hence even if we adopt a Humean fact/value distinction (and marry it up to the evaluative/non-evaluative in the way that has become so familiar to us) we have no grounds as yet for putting 'true' and 'means' anywhere but in the straightforwardly factual box. Consequently we have no handle here against scientific reductivist (e.g. functionalist or dispositionalist) accounts of thought and reason.

Many people share the suspicion that there is some flaw in such reductivist theories because the (causal and physicalist) materials they offer for building an account of mind and meaning could never be

---

[4] Kripke, *op. cit.* esp. pp. 86-95.
[5] Blackburn, *op. cit.* esp. pp. 286-7.

adequate to capture notions like 'reason', 'mistake' or 'mattering'. The gut feeling is perhaps that *meaning* is bound up with *pointfulness* in ways that the functionalist accounts are bound to leave out. But the moral of our remarks about truth is that whatever there is in this line of thought cannot be captured in the crude picture sketched above.

## II

Let us then turn to 'true' as a predicate of beliefs or judgements and ask what sense we can make here of seeing truth as a Good. Some philosophers speak of truth being the 'aim' or the 'internal goal' of judgement as though the formation of belief was an action that by definition required one to pursue truth as one performed it.[6] But this is a strange picture of things, if only because judging is not something that we do at will and hence not something the doing of which could be said to have any aim or goal.

So if we are to maintain that 'true' is evaluative as applied to beliefs we have to find some other active form of behaviour in the choice of which we express our positive opinion of truth. Inquiry is of course the obvious candidate here. We would express our regard for truth by choosing methods of enquiry which we regarded as likely to yield truth, by willingness to take pains and expend effort in order to have a better chance of arriving at true beliefs.

But when we take such pains and expend such effort (as, clearly, we often do) is it concern with truth itself which motivates us or is it concern with something else? It is obvious that, sometimes, it is something else. Consider for example a bank robber who has the project of burrowing through from the sewers to the bank vault. For him, information on the exact run of the pipes and the location of the vault is extremely important. He values true belief on these matters and takes steps to secure it. But his concern is that he should think that the vault is in such and such a place if it is in that place and should not think so if it is not. There is no interest of his that has to do with trust or truth in general, no interest of his that has to do with truth in any sense over and above that which a redundancy theorist might allow.

---

[6] B.A.. Williams, 'Deciding to Believe', *Problems of the Self*, Cambridge University Press 1973. S. Blackburn, *Spreading the Word*, Clarendon Press 1984, p. 231.

This observation does not support the view that there is no substantive account of truth to be given beyond the redundancy view. Let us suppose that there is such an account, say, a correspondence one: the world contains an immense number of facts, our minds contain a large (if not equally large) number of beliefs and a given belief is true if there exists a fact to which it corresponds. (If you do not like this account of truth, slot in whatever your favourite candidate is.) Given this it seems possible for someone to have an interest precisely in the promotion of as many correspondences as possible and in the elimination of facts and beliefs without suitable partners. It would be a project of the same shape as that of, say, providing a suitable gnome for every garden and a garden for every gnome, gnomes and gardens just not being at their best unless paired off. Or compare the aspirations of the dedicated match maker. But this promotion of correspondence, even if intelligible as a project (which we shall see reason to doubt), is neither here nor there as far as the bank robber is concerned.

The point I wish to stress here does not have to do essentially with the fact that the bank robber regards the acquisition of accurate information about the vaults and sewers merely as a means to the end of getting money. This may well be so in his case. And if there was another easier way of getting money perhaps he would pursue that. But there are other cases of a different kind which exhibit my central concern without sharing this feature. Suppose, for example, that I look in a bag to count how many apples there are, with a view to seeing whether I have enough to give one to each child. Imagine that (at least part of) my concern is non-self-interested, namely to give the children something they will enjoy if doing so is compatible with fairness. Carrying out this aim is only possible if I come to have a true belief about how many apples there are in the bag. I need to think that there are five if there are five, six if there are six and so forth. Having whichever of these beliefs is correct is not a mere means to my end (at least in some plausible ways of fleshing out the example). My end is conscious action by me of a certain character and my possession of an accurate conception of what I am doing is intrinsic to that action. One cannot act justly or kindly while sleepwalking or under hypnotic control by another, just as one cannot then get married or make a valid will. (There are interesting questions about what exactly would be lost and to whom if I were to distribute the apples under the hypnotic control of another. But that something would be lost seems clear.) So we have here a case of a consciously performed just and beneficent action which is not a contingent means to some other end but an end of itself. And my accurate conception of the number of

apples in the bag is an essential component of this end. Hence it is valued for itself, and not, as in the bank robber case, merely as a means. But in saying this we are conceding nothing to the view that correspondence is a good thing in itself or to the idea that truth exists as an independent value which is, at least potentially, in conflict with justice, beneficence and so forth. What matters is that I should think that there are five apples in the bag if there are five and so forth. But this matters only because it is a necessary element in a just and beneficent project.

Can we generalise the moral of these cases? To do so would not, as the second example makes clear, be to claim that truth is only valuable as a means, nor yet to support some form of egotism or hedonism. But it might be thought to involve the idea that the value of any belief was bound up with the value of some practical project of which it was an element. Hence the existence of the things we label 'the disinterested pursuit of truth' or 'the pure thirst for knowledge' might be adduced as a difficulty. Do these things not show that some motivation of the gnome and garden or matchmaker variety is available, at least to the more high minded or theoretically interested of us? I want to suggest that this is a mistake and to offer a different account of such cases.

The view that we value truth 'in itself' (whether we give a correspondence or any other substantive account of it) has the implication that any instance of truth is, merely in virtue of being true, worth having. (Compare: any specimen of the amusing is, *qua* amusing, at least a little bit worth paying attention to; any specimen of the painful is, *qua* painful, worth avoiding.)

Suppose then that I see someone walking slowly down the street every morning writing things in a notebook. I am told he is learning the numbers of the cars parked there. Many accounts of why this knowledge is worth having might be given – police enquiries, traffic surveys, snooping etc. But suppose that none of these is forthcoming. I am told instead that the person with the notebook is aware that there are some truths to be gathered about the car numbers and having a great thirst for knowledge has set out to learn them. We cannot make much of this. And there are innumerable similar cases of worthless potential knowledge which we are inclined to forget when we contemplate the 'disinterested pursuit of knowledge' in the case of historical or cosmological research.

One might argue that no item of knowledge is totally worthless because it might turn out to be important. Even though he did not know it at the time the car numbers that he notes later provide the vital evidence at a murder enquiry or in support of some fascinating sociological theory. But this move fails for two reasons.

The first is that it tries to move from 'If such and such then it would be important to know so and so' to 'It actually is important to know so and so, even if only to a small extent'. However the conditional premiss only supports the conclusion if supplemented by another – namely 'It is likely, or at least possible, that such and such' – which will license the detachment. But why should we always be willing to add this extra premiss? To do so would be in effect to allow that we can never know things of a certain character, e.g. that a certain piece of information will never be required in a murder investigation or any similar important inquiry. But I think I do know this about at least some bits of information. For example, I would claim to know that information about whether I put my left sock or my right sock on first this morning is of no interest and never will be of any interest. This is not to say that I cannot tell stories about conceivable development in which it does turn out important. But I claim equally to know that these conceivable events will not occur. We are here on familiar ground with problems about the relation of knowledge and possibility. But it is clear that only some controversial form of scepticism could support the conclusion I deny.

But there is a second reason why, even if we do concede some general form of scepticism, the case will still not do what is required of it. What we are debating is whether 'It is true' or 'It is a piece of knowledge' *just by itself* demonstrates the worth of the item in question. And the appeal to some condition upon which the item would have worth has tacitly conceded that it does not and has handed over the justificatory role to some value or purpose which would be in play if the condition were fulfilled.

Another attempt to keep going the idea of the value of truth in itself might propose that we find the car number case so bizarre because in order for action to be intelligible it is required not only that all truth be valuable but also the highly unlikely condition that the person in question has nothing better to do with his time than to learn car numbers. Suppose, however, that we imagine away the possibility of all more worthwhile pursuits, e.g. by locking our subject up in prison and depriving him of job, friends, exercise, books etc.. Do we not find then

that the tiny pull of these little truths makes itself felt? We can certainly imagine him paying minute attention to every detail of his cell layout, carefully memorising the shape and size of every stone in the walls and so forth. But an alternative construal of this behaviour is surely available – namely that it is better to pretend to have a purpose than to go mad. We can equally imagine him passing his time playing solitaire. Does this show that having one marble by itself in the middle of the board is valuable in itself? If it were, would he not achieve it more simply by tipping all the marbles out and putting one back in the centre?

The moral I would like to draw from the discussion is this: if someone claims that information on a certain topic would be a good thing one can always ask 'Why do you want to know about that?' An intelligible answer will have to say something about *that particular subject matter*. It cannot simply point back to the fact that the item in question would be a specimen of true belief. But to say that an answer must be forthcoming is not to say that the form of the answer must involve reference to some practical project in immediate or distant contemplation. To put it crudely, I am not arguing that only applied research is defensible. An intelligible answer would be that the subject matter in question was fascinating in itself and therefore worth investigating.

What is it for a subject matter to be worth investigating? Suppose we ask 'What's so interesting about that?' of some particular topic. Various responses are possible. One would be to point to the practical usefulness of the knowledge in question. But leaving that on one side, there seem at least two more available. The first would be to point out direct epistemological links between this proposed investigation and enquiries already afoot. ('If we could find out… then we could use that to determine whether…') The second would be to draw attention to analogies between the proposed investigation and ones already acknowledged as interesting. ('You know those fascinating results about…? We're trying to find out the same sort of thing about…')

But both of these responses presuppose the existence of some disinterested curiosity in the person to whom the answer is being given. What are we to say to a person whose question is about the point of any such thing? The likelihood is that discussion on this will run parallel to discussion of altruism. In both cases we have *prima facie* outward looking and non-self-concerned human endeavours, and in both cases the standing temptation is to try to provide justification for them which

undercuts this appearance, e.g. by appealing to some advantage to the self which is secured by their pursuit. It is true that a life without these concerns is impoverished and lacking in certain sources of profound satisfaction. This is one reason why parents try to promote such interests in their children. But arguably this truth is not one which could either persuade the egoist to acquire non-self-concerned interests or provide non-egoists with reason for their choice of life. This is because the very appreciation of what the impoverishment consists in requires that one already accept that the existence of other people and of the enormous and intricate world around us provide us with reasons for non-egoistic projects. The supposed justification is intelligible only to someone for whom it is redundant.

Now suppose that this controversial (and here inadequately supported) view were correct. Does that do anything to reinstate the idea of truth as a value in itself? It does not because the subject matter still remains of central importance to us in making intelligible any offered instance of pure curiosity.

The moral of this section could be summarised thus: there is no goddess, Truth, of whom academics and researchers can regard themselves as priests or devotees and whose service must be accepted as some justification for any endeavour. Rather there are as many different projects as there are different subjects of enquiry and each one of them will have its own justification in terms of usefulness or intrinsic interest.

As a corollary to this we might remark that academics and researchers are not burdened with some requirement of 'intellectual integrity' or 'absolute commitment to the truth' over and above what is required of anyone. To put the point more usefully the other way round, no one is let off the demands of intellectual integrity or commitment to truth because he or she is not at that moment engaged in academic research. The bank robber who has some personal stake in the view that the sewer runs in such and such a direction, fails to take adequate account of the evidence to the contrary and thus misleads his associates or himself, has lost the intellectual integrity he should have preserved just as much as the scientific researcher who unacceptably massages his data. Intellectual integrity is simply the virtue of theoretical reason whose counterpart on the practical side is courage. Both of them are requirements of any self-conscious rationality, i.e. of the ability to get into focus what one is about and to do what one sees to be demanded by

it. And as with courage intellectual integrity is at the service of bad projects as well as good.

*Jane Heal's*
*'The Disinterested Search for Truth'*

JULIAN DODD

# JANE HEAL'S 'THE DISINTERESTED SEARCH FOR TRUTH'

## JULIAN DODD

1. THE idea that truth is a value has an illustrious history. Frege famously claimed truth to be the aim of the sciences, placing 'true' into the evaluative bag along with 'good' and 'beautiful';[1] and in this he has been followed by countless others. Truth, it is generally supposed, is a norm: specifically, it is the aim of belief and the goal of enquiry.[2]

To focus our subsequent discussion somewhat, let us follow Krister Bykvist and Anandi Hattiangadi (2007: 277) in provisionally representing the claim that belief aims at truth as follows:

> (1) For any $S$, $p$: $S$ ought to (believe that $p$) if and only if $<p>$ is true.[3, 4]

Some philosophers doubt the coherence of quantification into sentence position.[5] Presumably, they will be happy enough if we rewrite (1) along the following lines:

> (2) For any $S$, $x$: $S$ ought to (believe $x$) if and only if $x$ is true.

In (1) – or, if we prefer, (2) – we have a tolerably clear initial interpretation of what is meant by the claim that belief aims at truth: a sense in which truth is supposed to be a Good Thing. The conditionals from right to left, and then left right, across the biconditional are supposed to accommodate, respectively, the two parts of what Christian Piller has described as '[William] James's Insight': namely, that we should both know the truth and avoid error (Piller 2009: 193-4).

---

[1] 'The word "true" indicates the aim of logic as does "beautiful" that of aesthetics or "good" that of ethics. All sciences have truth as their goal; but logic is concerned with it in a quite different way from this. It has much the same relation to truth as physics does to weight or heat. To discover truth is the task of all sciences; it falls to logic to discern the laws of truth' (Frege 1918: 17).
[2] Michael Lynch (2009: 10-13; 111-14) offers a particularly trenchant defence of these claims.
[3] Like Paul Horwich (1998: 10), I write '$<p>$' for 'the proposition that $p$'.
[4] (1), in its biconditional formulation, would seem to be the standard interpretation of the thesis that truth is the aim of belief. It is endorsed by Marian David (2001: 152), Alan Gibbard (2005: 338-9), and Lynch (2009: 10), to name but three.
[5] Wolfgang Künne (2003: 356-65), to my mind convincingly, rebuts such Quinean worries.

2. The key thesis of Jane Heal's engaging and compellingly written paper is that truth is not a value in this or any other sense. For according to Heal, 'the transparency of truth allows whatever valuable features there are in a situation or project to shine through but does not itself contribute anything of substantive value' (Heal 1987/8: 97). Heal does not herself state what the thesis of truth's transparency consists in, but this is easily rectified on her behalf. The transparency of truth consists in its possession of the following property: if one possesses the concept of truth, then to believe that, assert that, or enquire whether it is true that *p* is just to believe that, assert that, or enquire whether *p* (and conversely). For instance, for me to believe that

(3) <Eleanor is late> is true

is just for me to believe that

(4) Eleanor is late.

As Simon Blackburn puts it, 'it is as though you can always look through "it is true that" to identify the content judged, inquired after, and so on, as if the reference to truth was not there' (Blackburn 1984: 227).

It is a familiar point that truth's transparency enables 'true' to function as a handy device for making indirect and compendious endorsements of assertions, as in

(5) What Eleanor just said is true

and

(6) Everything Susan said is true,

respectively. The truth predicate's use as such a device in (5) and (6) is to be explained in terms of the predicate's function in the basic case: that is, in its explicit application to a proposition, as in (3). And the basic case is one in which 'true' is transparent: *i.e.* in which predicating 'true' of a proposition simply accomplishes what we can do by just asserting the proposition in question.

Now let us return to (1) and (2). Since 'true' is transparent, and since (1) sees 'true' predicated of an explicitly presented proposition, it follows that we can remove all mention of truth from (1), rewriting it as

(1*) For any *S*, *p*: *S* ought to (believe that *p*) if and only if *p*.

Consequently, as long as we are happy to use quantification into sentence position in our statement of the norm of belief, it will follow that this norm 'can be better [in this case, more simply] described in another way which does not involve the mention of truth' (Heal 1987/8: 97). Of course, 'true' cannot be likewise removed from (2); but this does nothing to show that the truth predicate is a contributor of evaluative content to this sentence, since (2) sees 'true' functioning as a device for facilitating compendious endorsements of assertions, and this is a feature the predicate has purely by virtue of being transparent.

What this reveals is that merely granting the correctness of (1) and (2) does not commit us to the idea that these sentences describe a norm of *truth*. Since the use of 'true' in both (1) and (2) can be accounted for on the basis of its possession of the transparency property, the predicate's occurrence in these sentences does not demonstrate that *it itself* expresses a value. Equally, however, it has not yet been shown that 'true' does *not* have evaluative content. For it is important to remember that 'true' can express a substantive – perhaps normative – property *and* be transparent (Wright 1992: 15). So how could Heal's claim that truth is non-normative be established?

3. Two strategies for substantiating Heal's key claim would seem to be in the offing. The first such strategy is to adopt some species or other of *deflationism* about truth. Deflationism, as I understand it, is the doctrine that all there is to the property of truth is its being transparent (McGrath 2002: 308). If deflationism is assumed, then truth is not a normative property and (1) and (2) cannot be taken to display normative facts about truth at all. On the contrary, for the deflationist, (1) and (2) merely illustrate the transparency of 'true' and, in the case of (2), its use to facilitate quantification over its primary bearers. Since (1) and (2) are universally quantified claims, both (1) and (2) state that believing is subject to the following normative constraint: roughly, one ought to believe that things are a certain way just in case that is how they are. For the deflationist, this is not a norm of truth, strictly speaking; it is just a norm whose colloquial statement in English – namely, 'One ought to believe something just in case it is true' – sees us use 'true' merely in order to put it into words.

Heal, however, does not follow this strategy. She leaves it open that there is more to truth than its possession of the transparency property (Heal 1987/8: 103). So how does she argue against the thesis that truth

is evaluative? By arguing that truth, even if a substantial property, is not evaluative in nature. Let us look a little closer.

4. As we have just seen, the deflationist about truth need not deny that there is a general norm of belief captured by (1) and (2). She can rewrite (1) as (1*), treat (2) as merely illustrating how 'true' facilitates generalisation with respect to sentence positions, and then claim that the norm involved is not a norm of truth. Heal's strategy, by contrast, would appear to be that of denying the existence of such a general norm of belief, on the grounds that (1*) and (2) are *false*. According to her, no such general norm exists, whether or not we classify it as a norm of *truth*.

To see why Heal thinks this, we need to look at her discussion of the cases that seem most strongly to indicate that (1*) and (2) are true: namely, instances of what we tend to describe as 'the disinterested pursuit of truth' (Heal 1987/8: 104). These are cases in which we value gaining true belief on a certain subject matter, but in which the value of the belief is not bound up with the value of any kind of practical project. Think, for example, of a pure mathematician's desire to discover whether some abstruse mathematical theorem – a theorem with no practical implications – is true. If asked to justify her enquiry, our mathematician will reply merely by expressing her disinterested curiosity in whether *p*: that is, her conviction that knowing whether *p* is valuable for its own sake. Does the existence of examples such as this not suggest that, 'at least to the more high minded or theoretically interested of us' (Heal 1987/8: 105), truth *in itself* is valuable: *i.e.* that (1*) is correct?

Heal thinks not. As she sees things, our pure mathematician's disinterested curiosity can be rendered intelligible only if we say something about the *subject matter* that so preoccupies her; and this indicates, not that it is truth *in itself* that is her goal, but that what she values is the truth on *that particular question*, whatever its details may be. 'There is', says Heal,

> no goddess, Truth, of whom academics and researchers can regard themselves as priests or devotees and whose service must be accepted as some justification for any endeavour. Rather there are as many different projects as there are subjects of enquiry and each one of them will have its own justification in terms of usefulness or intrinsic interest. (Heal 1987/8: 108)

In other words, there is no single norm governing belief, but a series of epistemic obligations, each specific to its subject matter: a distinct such obligation for each distinct proposition that can take the place of '*p*' in a true proposition of the form: <*S* ought to (believe that *p*) if and only if *p*>.


5. Heal's position is clear. Belief, as such, is not governed by a norm of truth. It remains to be seen, though, what Heal's *grounds* are for denying (1*): the universally quantified statement ascribing a general norm for belief. Is it not the case that belief *in general* aims at truth? Not according to Heal, and here is her reasoning. The idea that truth *in general* – that is, truth *in itself* – is a value of the kind specified by (1*) 'has the implication that any instance of truth is, merely in virtue of being true, worth having' (Heal 1987/8: 105). But as Heal points out, this result is extremely implausible, since it amounts to the thesis that we ought to believe even utterly trivial or uninteresting truths. There are, she observes, innumerable cases of 'worthless potential knowledge' (Heal 1987/8: 105).

I think Heal is right to deny that every truth is worth knowing; and, *pace* Michael Lynch (2004: 186-7), I doubt whether the counter-intuitiveness attaching to the claim that we ought to believe every truth can be dissolved by adding a '*prima facie*' clause to it. I surely do not have even a *prima facie* epistemic obligation to believe the multitude of trivial, mundane and uninteresting true propositions that come my way as I go about my daily business (David 2005: 297-8). As I idly read my morning newspaper, I cannot help but notice that the page has a small tear on its left-hand edge. Was there really a *pro tanto* obligation on me to believe this? A positive answer to this question looks a little far-fetched.

But in fact, Heal does not do justice to the full *degree* of the implausibility – one might say the *absurdity* (Engel 2002: 128) – attaching to the idea that we should believe every truth. For there are two convincing reasons for thinking that it is *metaphysically impossible* to believe everything that is true. First, there are infinitely many propositions. Second, given that any conjunction of atomic truths is itself a truth, there will be certain truths that are too complex for us to believe (Bykvist and Hattiangadi 2007: 279). Since 'ought' implies 'can', I cannot be obliged – even *prima facie* obliged – to believe every truth.

Heal is thus most certainly right to say that we are under no epistemic obligation to believe every truth. My worry, however, is that

she is wrong to think that she has thereby established that belief is not governed by *some kind* of norm of truth (as we might put it). For as Bernard Williams has explained, it is gratuitous to presume, as Heal does explicitly, that our valuing truth *in itself* entails that *every* instance of truth is worth having. This way of understanding 'in itself', 'as such' or 'for its own sake' is, Williams says, unreasonable. And he makes his point by drawing an illuminating analogy with the way in which we might describe someone's love of music. '"He loves music as such" or … "for its own sake"', Williams continues, 'does not imply that there is no music he does not want to hear' (Williams 2002: 286).

This strikes me as conclusive; and what it suggests is that there might be ways of preserving the intuitions behind the claim that belief aims at truth without committing ourselves to the consequence that Heal so rightly deprecates. Two such suggestions for amending (1*) have been prominent in the literature. Paul Boghossian suggests that we abandon the right-to-left conditional in (1*), formulating the claim that belief aims at truth as

> (1**) For any *S*, *p*: *S* ought to (believe that *p*) only if *p*.
> (Boghossian 2008: 100)

Boghossian believes that the gist of the norm for belief is that we should believe only what is true (*i.e.* that we should not believe falsehoods). Ralph Wedgwood, by contrast, preserves a right-to-left conditional by restricting the norm's scope to propositions that the believer has considered. According to Wedgwood (2002: 273), the norm of belief is best formulated as

> (1***) For any *S*, *p*: if *S* considers whether *p*, then S ought to (believe that *p*) if and only if *p*.

Needless to say, both of these attempts to reformulate the claim that belief aims at truth have been challenged. In particular, we are entitled to ask how (1**) can do justice to the idea – apparently implicit in the idea that belief aims at truth – that we should uncover truth, not merely avoid falsehood.[6] But I shall finish by just saying this: whether or not we want to deny that the aim of belief is, strictly speaking, truth, there might be more in the idea of such a norm than Heal thinks.[7]

---

[6] For a version of this criticism of (1**), as well as an interesting objection to (1***), see Bykvist and Hattiangadi 2007: 280-3.
[7] In private correspondence, Heal herself wonders whether any defensible attempt to restrict the claim of a general norm of truth – along the lines of (1**) or (1***) – will

REFERENCES

Blackburn, S. 1984. *Spreading the Word*. Oxford: Oxford University Press.

Boghossian, P. 2008. 'Is Meaning Normative?'. In his *Content and Justification: Philosophical Papers*. Oxford: Oxford University Press.

Bykvist, K. and Hattiangadi, A. 2007. 'Does Thought Imply Ought?'. *Analysis* 67: 277-85.

David, M. 2001. 'Truth as the Epistemic Goal'. In M. Steup (ed.), *Knowledge, Truth and Duty*. Oxford: Oxford University Press, 2001.

David, M. 2005. 'On "Truth is Good"'. *Philosophical Books*

Engel, P. 2002. *Truth*. Chesham: Acumen Press.

Frege, G. 1918. 'The Thought'. In P.F. Strawson (ed.), *Philosophical Logic*. Oxford: Oxford University Press, 1967: 17 38.

Gibbard, A. 2005. 'Truth and Correct Belief'. *Philosophical Issues* 15: 338-50.

Heal, J. 1987/8. 'The Disinterested Search for Truth'. *Proceedings of the Aristotelian Society* 88: 97-108.

Horwich, P. 1998. *Truth*, 2nd edn. Oxford: Oxford University Press.

Horwich, P. 2006. 'The Value of Truth'. *Nous* 40: 347-60.

Lynch, M. 2004. *True to Life: Why Truth Matters*. Cambridge, MA: MIT Press.

Lynch, M. 2009. *Truth as One and Many*. Cambridge, MA: MIT Press.

Künne, W. 2003. *Conceptions of Truth*. Oxford: Oxford University Press.

end up restating her main contention, *i.e.* that the value of any situation comes from its specific content and the role that knowing it would have in our lives. Specifically, she suspects that restrictions that are general or formal in some way (such as (1***)) will be counter-examplable by cases involving trivial, useless truths, whilst attempts to cite something more substantive in the restricting material will end up reiterating her particularist claim. I'm not sure about this, but I thank her for her characteristically astute comments.

McGrath, M. 2005. 'Lynch on the Value of Truth'. *Philosophical Books* 46: 302-10.

Piller, C. 2009. 'Desiring the Truth and Nothing But the Truth'. *Nous* 43: 193-213.

Wedgwood, R. 'The Aim of Belief'. *Philosophical Perspectives* 16: 267-97.

Williams, B. 2002. *Truth and Truthfulness*. Princeton, NJ: Princeton University Press.

Wright, C. 1992. *Truth and Objectivity*. Cambridge, Mass: Harvard University Press.

JENNIFER HORNSBY & GILA SHER

# BIOGRAPHY

Jennifer Hornsby has been Professor of Philosophy at Birkbeck, University of London since 1995, and has been Co-Director of the Rational Agency section of the Centre for the Study of Mind in Nature since its inception in 2007. She is an Emeritus Fellow of Corpus Christi College where she was a Tutorial Fellow from 1978 until 1994. She is a Fellow of the Norwegian Academy of Science and Letters. She has published work in philosophy of mind and action, philosophy of language, and on questions about free speech as they relate, among other things, to pornography. Jennifer was president of the Aristotelian Society from 1996 to 1997.

"The Disinterested Search for Truth" was originally published in the *Proceedings of the Aristotelian Society, Volume XCVII* (1997).

Gila Sher is a professor of philosophy at the University of California, San Diego. Her publications include *The Bounds of Logic: A Generalized Viewpoint* (1991), "Did Tarski Commit 'Tarski's Fallacy'?" (1996), "Is There a Place for Philosophy in Quine's Theory?" (1999), "In Search of a Substantive Theory of Truth" (2004), "Epistemic Friction: Reflections on Knowledge, Truth, and Logic" (2010), "Is Logic in the Mind or in the World?" (2011), "Forms of Correspondence: The Intricate Route from Thought to Reality" (Forthcoming), and "The Foundational Problem of Logic" (Forthcoming). She is an Editor in Chief of *Synthese* and a Consulting Editor of the *Journal of Philosophy*.

*Truth: The Identity Theory*

JENNIFER HORNSBY

# TRUTH: THE IDENTITY THEORY

## JENNIFER HORNSBY

I want to promote what I shall call (unoriginally, and for the sake of its having a name[1]) 'the identity theory of truth'. I suggest that other accounts put forward as theories of truth are genuine rivals to it, but are unacceptable.

A certain conception of *thinkables* belongs with the identity theory's conception of *truth*. I introduce these conceptions in Part I, by reference to John McDowell's *Mind and World*; and I show why they have a place in an identity theory, which I introduce by reference to Frege. In Part II, I elaborate on the conception of thinkables, with a view to demonstrating that the identity theory's conception of truth is defensible. Part III is concerned with the theory's relation to some recent work on the concept of truth: I hope to show that the identity theorist not only has a defensible conception of truth, but also, in the present state of play, has appropriate ambitions.

<center>I</center>

1.1 McDowell introduced the notion of a think*able* in order to fend off a particular objection to the following claim (1994, p. 27).

> [T]here is no ontological gap between the sort of thing one can... think, and the sort of thing that can be the case. When one thinks truly, what one thinks *is* what is the case.... [T]here is no gap between thought, as such, and the world.

Someone who objects to this supposes that, by denying any gap between thought and the world, one commits oneself to a sort of idealism. But such an objector confuses people's thinkings of things with the contents of their thoughts. If one says that there is no ontological gap between thoughts and what is the case, meaning by 'thoughts' cognitive activity on the part of beings such as ourselves, then one is indeed committed to a sort of idealism: one has to allow that nothing would be the case unless there were cognitive activity – that there could not be a mindless world. But someone who means by 'thoughts' the contents of such activity, and

---

[1] For 'the identity theory' in recent and contemporary philosophy, see Candlish 1995.

who denies a gap between thoughts and what is the case, suggests only that what someone thinks can be the case.

> [T]o say that there is no gap between thought, as such, and the world, is just to dress up a truism in high-flown language. All the point comes to is that one can think, for instance, *that spring has begun*, and that the very same thing, that spring has begun, can be the case. That is truistic, and it cannot embody something metaphysically contentious....

In order to avoid the ambiguity in 'thought' which would be exploited if a metaphysically contentious idealism were reached, McDowell suggests using the word 'thinkables' for what may be thought. My policy here will be to use the word 'thinkable' generally, in place of any of the more familiar 'content', 'proposition' or 'Thought'. Further reasons for this choice of word will show up in due course.

McDowell's demonstration that his position avoids a simple idealism may strike some people as an inadequate defence. I think that it can help to defend it to locate it by reference to debates about truth. One may view the quotations from McDowell as encouraging an identity theory of truth[2]. This says that true thinkables are the same as facts. True thinkables then make up the world of which McDowell speaks when he dresses up a truism. The world is 'everything that is the case', or 'a constellation of facts', as McDowell puts it, following Wittgenstein.

1.2 The identity theory is encapsulated in the simple statement that true thinkables are the same as facts. But it may be wondered how that statement could amount to a *theory* of truth: 'If someone asks what truth is, and receives an answer which helps itself to the idea of a fact, then hasn't she been taken round a very small circle?' Yes. But the simple statement on its own is not supposed to tell us anything illuminating. A conception of truth can be drawn out from an elaboration of what the simple statement can remind us of. And, as we shall see, the conception can be set apart from the conceptions of other accounts that go by the name of theories of truth.

The identity theory is not vacuous. It cannot be vacuous because it takes a stand on what the bearers of truth are, calling them thinkables. This is not an uncontentious stand. For there are philosophers who have

---

[2] I do not say that McDowell himself would see a point in viewing them thus.

told us that the notion of proposition (and thus of thinkable) is so dubious that we should take the truth-bearers to be sentences.[3] The identity theory proceeds without such doubts, taking it for granted that we can make adequate sense of what is meant when someone says, for instance, 'She told me something that isn't true'.[4] And the identity theory not only asks us to understand such 'something's in appreciating where truth is applicable, but it also asks us to understand such 'something's in saying what truth's applicability consists in. Certainly there is no illumination at the point at which the word 'fact' is resorted to in order to say what this applicability consists in. But the identity theory makes definite commitments nonetheless.[5]

1.3 Whether or not its title to be a theory can be made out, it may be unclear why the word 'identity' belongs in it. What could be the point in

---

[3] The doubts are induced by Quine's attack on propositions, which I touch on below (at §11.2, and see nn. 6 & 10).

    I think that someone who had never encountered logic or semantics might have encountered predications of truth to thinkables without encountering predications of truth to sentences; and the question *what truth* is surely concerns a concept which might feature in a language about which logicians and semanticists had never had anything to say. At a minimum, then, a philosopher who takes *truth* primarily as a property of sentences must say something about what appear to be its predication to thinkables. Although I accord priority to thinkables' truth here, I acknowledge that, when returning answers to particular philosophical questions, the application of 'true' to sentences is indispensable: see below, Part II. I acknowledge also that what appear to be predications of truth to thinkables may be treated as no such thing, as in the prosentential theory (see nn. 6 & 16). *Pro hac vice* I talk as if the surface appearances were sustainable.

[4] In saying that the identity theorist proceeds without doubts, I do not deny that hard work has to be done to give accounts of what appears to be talk about propositions/thinkables. An identity theory of truth evidently places constraints on such accounts. See e.g. Rumfitt 1993's account of the construction of propositions: Rumfitt's constructionalism goes hand in hand with a paratactic treatment of the logical form of sentences containing 'that'-clauses; but his kind of constructionalism might be entertained outside the context of such treatment.

[5] Candlish says, of what he calls a 'modest' identity theory, that it is 'completely uninteresting-trivial... precisely because it has no independent conception of a fact to give content to the identity claim' (1995, p. 107). Candlish here assesses the theory as if it had the ambitions of a definition. But what I call 'the identity theory' has no such ambitions (and indeed it acknowledges truth's indefinability: see below). The interest of the theory derives from what it can be seen, from what it says, to be opposed to philosophically. Candlish allows that an *im*modest ('robust') identity theory might be interesting: its interest could derive from its 'independent conception of facts', independent, that is, of the conception of thinkables, or truth-bearers. For my own part, I cannot see a point in thinking that such a theory deserves the name of *identity* theory. (Here I disagree with Julian Dodd 1995, from whom Candlish takes the robust/modest distinction. There is much about which Dodd and I agree, however: see our 1992.)

saying that true thinkables *are the same* as facts, rather than – more simply and apparently to the same effect – that true thinkables *are* facts?[6]

A familiar argument in Frege (1918) may help to show the point. It is an argument against the correspondence theory of truth. Frege introduces it with the words 'It might be supposed... that truth consists in the correspondence of a picture with what it depicts'. 'This is contradicted, however', he says, and then argues by *reductio* (pp. 18-19):

> A correspondence... can only be perfect if the corresponding things coincide and so just are not different things at all... [I]f the first did correspond perfectly with the second, they would coincide. But this is not at all what people intend when they define truth as the correspondence of an idea with something real. For in this case it is essential precisely that the reality shall be distinct from the idea. But then there can be no complete correspondence, no complete truth. So nothing at all would be true; for what is only half true is untrue.

---

[6] The introduction of 'identity' might seem to have the consequence of upping the ontological stakes (so that thinkables are to be treated as OBJECTS). That is not so. When we have understood, for example, 'She does it in one way, and he does it in another way', we have also made sense of 'They don't do it the *same* way' – but not at the expense of treating either things that are done or ways of doing them as OBJECTS. I think that hostility to propositions derives partly from Quine's assumption that all quantification is objectual or (in Quine's own sense) substitutional. This assumption has seemed to have the consequence that unless we give a Quinean substitutional account of these 'something's,we shall be forced to treat propositions as OBJECTS, in a sense of the term caught up with a particular understanding of singular reference. But Quine's assumption is not compulsory: see e.g. Davies 1981, Ch. VI, §3. Some of the interest of the prosentential theory of truth, defended in Grover 1992 and Brandom 1994, derives from the directness of its challenge to Quine's assumption.

The identity theory is not formulated in order to take a stand on the logical form of predications of truth. If taken to reveal logical form, it would take an erroneous stand – the one which is contradicted by Frege's remark that "'true" is not a relative term'. Comparison with Russell's Theory of Descriptions may be helpful here. In the analysis of 'the' provided by Russell, the word 'the' is not treated as the simple quantifier which, presumably, so far as logical form is concerned, it is. One point of giving the analysis which Russell's theory states is to show what is involved in seeing 'the' as a quantifier, and to show which quantifier it is. Something analogous goes on when 'identity' is introduced into an account of truth. Just as Russell's Theory can present the negative semantical claim that 'the' does not combine with predicates to form names, so the identity theory of truth can present its own negative metaphysical claims — claims such as emerge from seeing how the identity theory arises out of rejection of a correspondence theory.

One point of a formulation including 'same' might be to draw attention to the principles of distinctness of facts presupposed to the theory: those principles cannot allow a coarser grain to facts than to thinkables. (This means that it is not a target of the so-called slingshot argument; see Neale 1995.) A naive account of facts, attractive to those who seek facts in line with a correspondence conception, might incorporate the principle: Where $a = b$, '$Fb$' does not express a different fact from '$Fa$'. Such a principle, obviously, is at odds with the identity theorist's conception of facts. (In Neale's terms: 'the fact that ( ) = the fact that ( )' is – PSST.)

Putting this only slightly differently, we hear Frege saying: if truth were explicated in terms of any relation, it would have to be identity, since anything less than a candidate for truth's coincidence with a putatively corresponding thing would lead to the intolerable conclusion that there is no truth. Someone who takes herself to think that true thinkables correspond to the facts has it right, then, only if she actually means that any true thinkable is the same as some fact – which is what the identity theorist says.

Frege's argument has a sequel. This starts by showing how Frege thinks his opponent will respond. The opponent asks (p. 19):

> But can't it be laid down that truth exists where there is correspondence in a certain respect?

Here it is conceded that truth cannot be unspecified correspondence, so to speak. The problem with taking truth to be unspecified correspondence is that there can be correspondence in this respect, or that respect, or that other respect, so that there can be less or more correspondence according as there is correspondence in fewer or more respects; but there can't *in any analogous way* be more or less truth.[7] The opponent supposes that he can get out of this difficulty by picking on one respect of correspondence. To this Frege has a response.

> But in which [respect]? What would we then have to do to decide whether something were true? We should have to inquire whether an idea and a reality, perhaps, corresponded in a laid-down respect. And then we should have to confront a question of the same kind, and the game would begin again. So the attempt to explain truth as correspondence collapses.

---

[7] Frege pointed out that 'with every property of a thing is joined a property of a [thinkable], namely that of truth' (p. 20). For illustration, suppose that Fred is tall. Putting it in Frege's way, a property of Fred (being tall) is joined to a property of a thinkable: if Fred is indeed tall, then a true thinkable is put forward when Fred is said to be tall. But if this is correct, then it can seem that we should allow that *truth* can have any of the features which the property of *being tall* can have, so that if *being tall* admits of degrees (if x can be to some extent tall), then *truth* admits of degrees (it can be to some extent true that x is tall). But now it seems that Frege appreciates a characteristic of 'true' which ensures that,when treated as a predicate, it will seem to admit of degrees, if any does. This makes me think that when Frege invokes the claim that what is half-true is untrue, he is relying on the thought that any *relation* introduced to account for truth cannot be a relation which admits of degrees. And that is why I say that there cannot *in any analogous way* be more or less truth.

The idea now is that if there was something distinct from a thinkable such that establishing that some relation obtained between it and the thinkable was a way of getting to know whether the thinkable was true, then someone could be in the position of knowing what is known when the thinkable is known, yet of still not knowing whether it was true. But of course one could never be in that position: to discover whether $p$ is already to discover whether it is true that $p$.

This reveals a general difficulty about defining truth – the difficulty which shows up 'when we confront the same question again'.

> In a definition certain characteristics would have to be stated. And in application to any particular case the question would always arise whether it were true that the characteristics we represent.

'Consequently', Frege concludes, 'it is probable that the word 'true' is unique and indefinable' (p. 19).

When one follows Frege's argument through to this general conclusion, about the definability of truth, explicit opposition to the correspondence theory is lost: the correspondence theorist's definition fails to meet a constraint on any adequate definition; but it turns out not to be alone in that failure. Frege accordingly might be thought to have argued against an especially naive correspondence theory in the first instance, and then turned to opposing the whole idea of truth's definability. But there can be a point in thinking of Frege's initial argument as meant to show that a correspondence theory in particular – and *any* correspondence theory – is untenable. This is an argument which is sound only if the identity theory escapes its *reductio*. Its conclusion may be dressed up in high-flown language: there cannot be an ontological gap between thought ('an idea') and the world ('something real').

1.4 The identity theory, at any rate, is distinguishable from any correspondence theory. And the identity theory is worth considering to the extent to which correspondence theories are worth avoiding. I think that correspondence theories *need* to be avoided. I mean by this not merely that they are incorrect, but that people are apt to believe them.

It is common for philosophers to speak as if a correspondence theory of truth had no metaphysical import whatever. We are sometimes told that the idea of correspondence is recorded in a series of platitudes that any theorist of truth has to respect. Simon Blackburn has spoken of the phrase 'corresponds to the facts' as sometimes a piece of Pentagonese – a

paraphrase of 'is true' deployed with the purpose of saying something important sounding (1984, p. 255). But of course this is not all that has ever been read into the phrase. Someone who says 're-rendered it operational' for 'got it going again', may be criticized for needless portentousness, but not on other grounds; but when 'corresponds to the facts' gets in, the phrase's wordiness should not be the only source of doubt. Certainly there are glosses on 'is true' that are platitudinous: 'is a fact' is one such – the one that the identity theory singles out for attention. Perhaps it is also a platitude that true sentences say how things are. And this again is unobjectionable, so long as the 'things' in question are ordinary objects of reference: the true sentence 'that book is red'; for example, says something about how things are by saying how one of the things (sc. that book) is (sc. red).[8]

From the point of view introduced by the identity theory, it will be distinctive of correspondence theorists to seek items located outside the realm of thinkables, and outside the realm of ordinary objects of reference, but related, some of them, to whole thinkables. The idea is widespread, and it takes various guises. In the Russell of *An Inquiry into Meaning and Truth* (1940), the basic correspondents are percepts. Percepts can be 'surveyed but not defined'; utterances appropriately associated with them get their particular meanings from them; and propositions, the truth bearers, can be constructed out of percepts. In the Quine of *Philosophy of Logic*, the correspondents are cosmic distributions of particles. 'Two sentences agree in objective information, and so express the same proposition, when every cosmic distribution of particles over space-time that would make either sentence true would make the other true as well' (1970, p. 4). These very different candidates for things that make sentences true – percepts and particle distributions – reflect the very different obsessions of Russell and Quine, epistemological and cosmic. But what is common to their accounts, despite this vast difference, is a willingness to reconstruct thinkables from posited entities of a different sort, entities which make things true.

---

[8] This platitude points up the independence of thinking from what there is. Whether you want to know the book's colour, or to know something of what I think about the book, you have to think of something that is not sustained in existence by your thinking. But the thing to which you are then related (that book) is obviously not something which corresponds to a thinkable.

Davidson used to say that a relation like Tarskian *satisfaction* could provide the language-world links sought by a correspondence theorist of truth. But Davidson now regards this as a mistake (1990, p. 302). It must indeed be a mistake if opposition to correspondence theories can be combined with thought about mind-independent objects.

Percepts and particle distributions, then, are supposed to be items which we can specify independently of an account of thinkables, items which may confer truth upon a thinkable. When they are introduced, however, we cannot hold onto the truism that inspires the identity theory. The fact (as it is) that autumn has begun, if it were to be a cosmic distribution of particles, would not be the same as what I think when I think (truly) that autumn has begun.

It is evident now that the words 'corresponds with' do not have to be in play for an ontological gap between thought and the world to open up. This is something that we see in formulations used over the years by Michael Dummett and Crispin Wright in stating the semantic anti-realist's case. Their formulations often appear to invoke a conception of a truth-maker which will suit a correspondence theorist but which an identity theorist cannot allow.[9] Dummett asked 'If it were impossible to know the truth of some true statement, how could there be anything which *made* that statement true?'. Wright spoke of 'a truth-conferrer for a sentence': in the case where the truth of the sentence cannot be known, he said that this is something that 'the world fails to deliver up'. And he spoke of 'the states of affairs' that are in question when a sentence is undecidable as things that 'could not be encountered'. These ways of speaking give rise to an image of something with which a thinkable might have connected up, but a something which we are expected to think of the world as taking sole responsibility for. This is the image that an identity theory may help to rid us of. For when the conditions for the truth of a sentence are supplied by an identity theorist, nothing is brought in besides the thinkable that is expressed by the sentence itself. By introducing 'sources of truth', 'truth conferrers' and 'states of affairs', Dummett and Wright drive a wedge between what is demanded by a

---

[9] See Dummett 1976, at p. 61 in the version reprinted in Dummett 1993.

A different sort of illustration may be got from Frank Jackson, Graham Oppy and Michael Smith 1994. They argue for the compatibility of versions of non-cognitivism (in ethics, say) with minimalism about truth. They follow Michael Devitt in characterizing minimalism as holding that 'terms for truth and falsity are linguistic devices for talking about reality by appending the truth predicate'. Their claim then is that it might not be that any old sentence is such as to talk about reality: non-cognitivists, they say, 'precisely deny that (e.g.) ethical sentences talk about reality'. But someone who is opposed to correspondence theories in all their versions will not allow this 'talking about reality'. Suppose that Devitt had characterized minimalism by saying that truth and falsity are terms for going on talking while adding a word or two. Would Jackson et al. then have said 'Non-cognitivists precisely deny that (e.g.) ethical sentences are used in talking'?

This example may serve to show how easily ideas of correspondence get in through the back door.

thinkable and what is demanded by a thinkable that is true. The identity theorist leaves no room for any wedge.

Of course these remarks about Dummett and Wright do not get to grips with the position which was their concern. But they can illustrate a point – that philosophers' formulations are apt to create an outlook which is forsworn when an identity theory displaces a correspondence theory. I hope that they also suggest how the identity theory may displace forms of anti-realism more subtle than the crass idealism which results from equating thinkables with thinkings of them.

<div align="center">II</div>

2.1 It would be laborious to attempt to show that the identity theory is incompatible with all things irrealist. In order to show that it embodies nothing metaphysically contentious, I shall attempt only to reveal its actual compatibility with a perfectly common-sense realism.

McDowell's rebuttal of any simple idealism emphasizes the independence of thinkables from thinkings. One way to grasp this independence is to see that there are (so to speak) more thinkables than there are thinkings. I suspect that those who find the theory problematic are apt to suppose that it could be part of commonsense that there are (so to speak) more facts than there are true thinkables. If this is right about where the opposition lies, then further reflections on the identity theory, if they are to serve as a defence, must expand on the notion of a thinkable. By the identity theorist's lights, our grasp of the notion of a fact cannot exceed our grasp of the notion of a true thinkable. But someone who wishes to express doubts on that score might be helped by having it made apparent how generous the notion of a thinkable nonetheless is.

2.2 There can seem to be an immediate obstacle, however, to *any* account of thinkables – of the contents, the meaningful things that bear truth. Quine's attack on the Myth of the Museum is directed against the assumption that there could be things external to thought and meaning, lodged like exhibits in the mind, whose relations to other things could constitute the foundations of meaning (1960). The identity theorist agrees with Quine about the incoherence of the hope that intersubjective sameness of meaning might be explained in terms of relations with things external to thought and meaning.

From the identity theorist's point of view the correspondents of correspondence theories of truth play the same role as the exhibits in the museum of the mind: they are items located beyond the bounds of human play with concepts, in terms of which one is supposed to explain meaning. As Quine puts it, speaking himself of particle distributions, the item assigned to one sentence as a condition of *its* truth is the same as the item assigned to another sentence as a condition of its truth if and only if the two sentences have the same meaning. But such items as cosmic distributions of particles are in the same boat as items in the mind's museum according to the identity theory: neither can be used in the reconstruction of thinkables from something else.

If one countenances the cosmic items, but is led by the problems of the items in the mind's museum to think that ordinary talk of meaning is unsupported, then one may invoke a double standard. Quine tells us that a second class standard is appropriate so long as we are tolerant of such everyday psychological talk as involves any notion of a thinkable (1960, §45). But he said that we can, and in science we must, employ a first class standard; it is then that objective information, corresponding to (say) cosmic distributions of particles, can do duty for thinkables, Quine thinks. The upshot of this is hard to make coherent. For the view of everyday reports of people's psychological states which is required by Quine's lower standard for them is not a view that can be sustained by someone who takes herself (for instance) to seek the truth in some area. A person's being an enquirer of any sort requires that she be interpretable as aiming at gleaning the facts, and we have no conception of what that is excepting as we can think of her as more generally intelligible – as apt to perceive things, and to think them, and to draw conclusions. We cannot then be in a position to make statements about Quine's first-class reality but of refusing (according to the same standard) to make any statements which say, for instance, what people are doing when they are investigating that reality. The identity theory helps to make this difficulty with the Quinean picture vivid. The first-class standard was meant to be the standard of genuine facts; the second-class standard was to be invoked when the language of thinkables was used. But if any fact is the same as some true thinkable, then we cannot endorse facts and despise thinkables.

2.3 It can seem as though the identity theorist had nowhere to turn for an account of thinkables. At least there is nowhere to turn for an account besides an investigation of other predications to them – predications other than 'is true'. This brings me to further reasons (which I said I would come to) for using the term 'thinkable'.

'Thinkable' is a word for a sort of things to which a person can be related in various modes. I say that the Labour Party will win the next election. I have just said something (that Labour will win) which many now believe, which a good few hope, which John Major fears. The example then shows that thinkables can be beliefs, hopes and fears. They are called beliefs when thought of in connection with one psychological attitude towards them; they are called hopes or fears when thought of in connection with other attitudes. They are thought of as propositions when thought of as propounded.[10] A modal term, like 'thinkable', may serve to remind one of the variety of relations here: it is not only thought which relates to think*ables*, because a thinkable can be believed and hoped, for instance. (And just as we must not confuse a thinkable with a thinking, so we must not confuse a thinkable with someone's believing one, or with someone's hoping one.)

Besides '----- is true', then, there are predicates of thinkables, such as '----- is believed by Tony', '----- is hoped by members of the crowd'. Yet other predicates of thinkables show people as related to them by their speech acts: a statement, for instance, is what we call a thinkable when we think of it in connection with someone's making a statement. 'Thinkable' gives a word for what is truth-evaluable which is indifferent between the case where the evaluable thing is presented as the object of a state of a thinker's mind and the case where it is presented as having been put into words. But it is the linguistic expression of thinkables which we are bound to focus on, if we are to find anything of a systematic sort to say about them. One aim of theories of meaning is to show the significance of sentences as systematically dependent on properties of the words that make them up: theories of meaning, one might say, treat of thinkables' composition. The productivity of language, which can be revealed in its theory of meaning, then points towards another reason for using a modal notion, and speaking of think*ables*. Someone in possession of a theory of meaning for some language can say what was expressed in the use of any of the sentences on some list, composed from some stock of words; and is in a position to see that there are other things that would be expressed in the use of

---

[10] It seems worth remembering that propounding is a propositional attitude, and that Quinean hostility to propositions is hostility equally to beliefs (say). Because the opposition to certain abstract conceptions of thinkables has typically been directed against things called propositions, we find philosophers whose attitude towards beliefs and statements is one of acceptance, but towards propositions is one of rejection. (See e.g. David 1994, p. 12.) Of course it might be stipulated that the term 'proposition' is to mean what is meant by those who use the term illicitly. But short of making such a stipulation, it will be hard to justify an attitude of hostility peculiarly to propositions.

other sentences, not on the list, but composed only from words in the same stock. A theory of meaning, though its data are uses of actual sentences, is a theory which speaks to potential uses – to what would be said if some hitherto unused sentence of the language were used. There are actually unused sentences, which, just like the sentences we have given voice to or heard or read, express thinkables.

This suggests the place to look if we want to expand on the notion of a thinkable. We cannot postulate meanings in the mind or correspondents in the world. But we can look to the actual practices of language users. And we shall be reminded here of an idea first recommended by Donald Davidson – that we might put to work, as a theory of meaning of the language of some speakers, a definition of truth for the language which enables the interpretation of those speakers. Davidson's claim that a definition of truth for a language can serve as its theory of meaning depended in part on his thinking that Tarski had shown a way of displaying the recurrent significance of words – by treating words as having characteristics which affect the truth of sentences they come into.[11] In the present context, much of the importance of the idea of deploying such a definition of truth for a language is the view of predications of thinkables it affords. Where an account of a language's workings is interpretive of its speakers, it enables the theorist to give expression, in the case of any sentence in the language and any speaker of it, to the thinkable expressed by the speaker using that sentence. It thus gives the theorist the resources to say what speakers are doing when they use their language.[12]

An interpretive account of speakers is not narrowly linguistic. For speakers' productions of sentences cannot be seen as intelligible expressions of thinkables except as speakers are seen to have some

---

[11] Davidson 1967. I use '*definition* of truth' here as Davidson did there; and this allows me to avoid using 'theory of truth' ambiguously. (It seems impossible to avoid all possible ambiguity, however. Where a theory of truth [in the only sense of that phrase I use here] purports to give a definition, it purports to give a definition of truth; but of course what it purports to give is not the sort of language-relativized thing that Tarski showed one how to construct.) One makes no assumptions about Tarski's own intentions in saying that Tarski in fact showed us a way to construct a definition of truth for L that can be used to do something that a theory of meaning for L has to do. (Etchemendy 1988 has an understanding of Tarski's purpose which leads to a view of a definition of truth for a language which encourages the deflationist attitude discussed in §111 below.)

[12] I cannot here do more than take for granted a vast body of literature which shows the workability of definitions of truth for languages having natural languages' features. See, e.g., further papers in Davidson 1980. Davidson's idea has been endorsed by many others, of whom, in the present connection, McDowell should be mentioned; see, for example, his 1976.

purpose in producing the sentences. And any hypothesis about the purpose of a person who uses words on some occasion goes hand in hand not only with a hypothesis about the thinkable then expressed but also with hypotheses about her mental states about how *belief* and *desire* and the other attitudes relate her to thinkables – and with hypotheses also about the states of mind of audiences to her speech, and of all the others who use the language on other occasions.

The imaginary theorist, who compiles the facts about words that could put one in a position to understand foreign speakers, would be involved not only in making attributions to speakers of psychological attitudes and speech acts towards thinkables, but also, and inevitably, in taking a view of the truth of the thinkables to which speakers are then taken to be related. One cannot generally take a view about what someone's purposes are without having some view of which of those purpose are achieved; people intentionally do what they try to do to the extent that the beliefs which explain their doing what they do are true (are believings of true thinkables, that is). Of course the word 'true' does not have to be dragged in in order to see someone's taking an attitude towards a thinkable as working as it does. One can just as well say 'She believed that the 'plane took off at 9, and the plane took off at 9' as one can say 'She believes that the 'plane took off at 9 and that is true'. But insofar as an interpretive account requires more than the idea of people's relations to thinkables, and more than the idea of interconnections between those relations, it requires grasp of the distinction involved in assessments of thinkables as true or false. The view of thinkables that emerges, then, in trying to expand on the notion, is one in which some thinkables are taken to be (the same as) facts.

The study of interpretive accounts affords a distinctive perspective on the application of 'is true' to thinkables. 'True' can be treated as having a role alongside a variety of psychological predicates – but it is not itself treated as a psychological predicate, of course.[13]

---

[13] 13. Cp. Davidson 1990, p. 287: 'the concept of truth has essential connections with the concepts of belief and meaning'; and 'what Tarski has done for us is to show in detail how to describe the kind of pattern truth must make'. Davidson himself thinks that the empirical evidence we need in order to identify the pattern must avoid, in the first instance 'states with (as one says) a propositional object'. Davidson, then, would not be happy with the introduction of, 'as one says', propositional objects (i.e. thinkables) at the outset. This explains why his objections to Paul Horwich begin at an earlier point than my own do (see n.19). For his part, Davidson has a theory of verbal interpretation to elaborate: see 1990. To question the need for this would take me too

2.4 Discussions of coming to understand a foreign language sometimes assume its speakers to be more ignorant than the theorist: the facts at the theorist's disposal go beyond any of which the interpreted people are apprised. But this assumption is not essential to the idea of an interpretive account. Contemplating interpretive accounts shows the acceptability of a conception of potential uses of language expressive of thinkables outside one's ken, and some of which are facts.

One might think inductively here. Over the centuries, human knowledge, at least in some spheres, has expanded, and its expansion has been assisted by the introduction of new concepts, for instance in the formulation of scientific theories. If one believes that human knowledge will continue to expand, one is entitled to predict that thinkables which none of us here and now is capable of thinking will come to be known. One may envisage a theorist interpreting a language of the future: its speakers would think things, and the theorist, in coming to understand them, would learn from them. She could come to have access to facts, which in her present situation she is not even equipped to express.

Here one thinks of thinkables in connection with expanding knowledge. And it might then be supposed that the facts are to be circumscribed by reference to what is known by an ideal knower, at the limit, as it were, of an inductive series of more and more knowledgeable beings. But acceptance of unthought thinkables, some of which are facts, requires no such supposition. The supposition requires an understanding of the ideal situation for arriving at knowledge. And this can only be a situation in which all sources of error are eliminated or taken account of – a situation, that is to say, in which one is sure to believe what is *true*. Perhaps we can gesture towards such an ideal. But since we can explain it at best in terms of an antecedent notion of truth, the style of thinking used here to uncover a conception of facts can lend no support to an epistemic theory of truth.[14]

The conception of unthought thinkables elicited here does not depend upon any settled opinion about human ambitions or limitations, but only upon an idea of intelligible others from whom one could learn.

---

far afield. But I can try to state Davidson's view in my own terms: such a theory of verbal interpretation has to be understood from the standpoint of someone contemplating an interpretive account in order that such contemplation should ensure that a philosophically adequate conception of truth is elicited.

[14] Here I am thinking of, for example, the theory which seems to be endorsed in Putnam's 1981, which says that truth is an idealization of rational acceptability. In later writings (e.g. 1990), Putnam asks us to read the remarks he makes in supporting his Internal Realism as meant only to convey a picture, rather than as a theory of truth.

It evidently yields a generous conception of facts, to which an identity theorist is entitled. I hope, then, that the identity theory emerges as a defensible theory of truth, in keeping with our commonsensically realist view about the extent of facts independent of us.[15]

## III

3.1 However defensible the identity theory of truth might be made out to be, it might seem that there is an alternative to it – in the minimal theory of truth.

The minimal theory is advanced by Paul Horwich (1990, pp. 6-7).

> [I]t contains no more than what is expressed by uncontroversial instances of the equivalence schema:

---

[15] The remarks of this section are intended to go further than those of McDowell (reported in §1.1) – further towards showing that it is not a difficulty for the identity theory that it circumscribes the world using the notion of a thinkable. Although offered in defence of the claim that an identity theorist has a commonsensically realist conception of facts, they are not offered as a defence of any 'Realism' meriting a capital 'R'. In defending his 'Internal Realism' (see n.14), Putnam's target was 'Metaphysical Realism', a doctrine which the identity theory is evidently also opposed to. Of course it is possible to think that a defence even of commonsense realism is required: Michael Dummett has long urged this. In his 1990, Dummett thinks of the 'tacit acquisition of the concept [of truth]' as involving 'a conceptual leap;... just because this is so, it is open to challenge' (p. 200, in 1993 reprinted version). The leap, Dummett says, is one 'we all [made] at an early stage in our acquisition of our mother tongues': it involves a transition from the 'justifiability condition of an assertion to the truth-condition of the statement asserted' (p. 198). Now Dummett's own understanding of the conceptual leap is shown in his speaking of the notion of justification as 'cruder' and of truth as 'more refined'. But Dummett's opponent may resist any picture of the concept of truth as got from something cruder – as if there were something which might be added to justifiability to get truth, so that the child at some stage had to acquire the added extra. (The identity theorist seems bound to resist this, since she cannot allow *truth's* applicability to be separated from *thinkability*.) Against Dummett, it may be said that the child who comes to belong to a community of speakers (a systematic account of whose uses of sentences deploys the concept of truth) is drawn into practices in which the concept already has a place. Evidently in saying this, one still does not supply the defence which Dummett seeks. But perhaps it helps to make it clear that one can reject Dummett's story about the acquisition of the concept of truth while acknowledging that truth is indeed in an obvious sense more demanding than justification.

> (E) It is true *that p* if and only if p.

Horwich calls this 'the deflationary point of view'. In advocating a minimal theory, he means us to think that those who have waxed philosophical about truth in the past have tried to say too much and overshot the mark. He believes that we are apt to have an erroneously inflated conception of truth.

There has been so much writing under the head of 'minimalism' and 'deflationism' that in order that something should be fixed, I shall use Horwich's position to define 'a minimalist theory'. Deflationism, on the other hand, I shall treat as an attitude towards truth which a minimalist theorist takes, but which is also taken by others – disquotationalists, and, it seems, Richard Rorty.[16] In an attempt to make out the identity theory's superiority to the minimal theory, I start by suggesting that, despite what they have in common, there has to be a genuine difference in their conceptions of truth. Then I suggest that to the extent that the minimal theorist wants to convey a deflationary message about truth, which is not already conveyed in the identity theorist's opposition to correspondence, the message has to be resisted.

3.2 One thing that the identity theorist and minimal theorist agree about shows up in connection with a point that Dummett once made (1958/9). Dummett famously said that an advocate of a minimal theory is ill-placed to tell us that *truth* can be used to explain *meaning*.[17] The point

---

[16] Disquotationalists differ from Horwich in taking the truth of *sentences* to be primary, so that I have taken a stand against their position already (see n.3 above). For criticisms of disquotationalism as such, see David 1994. Many of these criticisms have versions applicable to Horwich's theory: in connection with Horwich they would start from asking what is involved in the acceptance of propositions – which is the question that I press below.

For Rorty's deflationism, see n. 21.

The characterization of deflationism here is deliberately vague (it is meant to be as vague as the statements used to convey the deflationary message, see §111.3). But I should note that, with Horwich's minimalist theory used as the paradigm of a theory provoking the deflationist attitude, it is *not* a characteristic thesis of deflationism to deny that truth is a predicate. Brandom 1994 takes his treatment of '... is true' as a prosentence-forming operator to secure one of deflationism's characteristic theses. But I think that the identity theorist's opposition to the deflationist attitude that Horwich means to provoke might survive arguments about the correctness of pro-sententialism. (From Brandom's position, one would see these issues in a different light. I cannot speak to it here, but I make a further remark about it at n. 23 below.)

[17] Dummett was actually talking about the redundancy theory of truth. For the purposes of considering his argument, we may think of this as a species of minimalist theory. Dummett himself has called Horwich's theory (which Horwich calls 'minimal')

relies on two features of the minimal theorist's conception. On the one hand, *truth* is not a notion of substance, which can be utilized in accounting for something else; on the other hand, *meaningfulness* is simply taken for granted when one predicates truth. We find parallel features in the identity theorist's conception: On the one hand, facts are thinkables, so that there is nothing external to thought and meaning in terms of which truth can be understood; on the other hand, thinkables are what facts are, so that *thinkability* is presupposed when one talks about truth. These features of truth presented the obstacle that we initially saw to giving any account of thinkables (see §II.2). And they explain the lack of illumination which is felt at the point at which the identity theorist resorts to the word 'fact' (cp. n. 5).

To the apparent obstacle as it presents itself to the identity theorist, I have responded with the idea of an interpretive account of a group of language users. Taking the perspective that such an account provides, one does not purport to *explain* either meaning or truth. (Of course not: Dummett's point is well taken.) But one is in a position to elicit features of one's conception of truth. By turning to interpretive accounts, one can shed light simultaneously on truth and thinkability (where an account of linguistic meaning is situated in an account of thinkables). One does so by contemplating the use of definitions of truth which are such as to contain all the instances of the schema found in Tarski's Convention T (the analogue, as it were, for sentences, of Horwich's Schema (E)). But it is here that a real disagreement with the minimal theory is revealed. For the idea of an interpretive account is alien to the conception of truth that Horwich means us to take away from his theory. Horwich for his part defends not only a minimal theory of truth, but also minimalism about meaning; and he defends minimalism about meaning by doing his best to show that the concept of truth is not needed in the explication of meaning. It is supposed to be a consequence of a minimal theory, that 'true', being deflated and shown to have no substance, could not play a crucial part in something as interesting as an interpretive account is thought by its defenders to be.[18]

---

'redundancy'. (The name 'redundancy theory', however, is *sometimes* reserved for a theory in which predicate-hood is denied of 'is true', i.e. for a species of a kind of theory mentioned at nn. 16 & 23.)

[18] On this reading, Horwich's argument to deflationism about *meaning* is indirect. (Truth has no substance, so meaning doesn't either.) But Horwich sometimes defends minimalism about meaning more directly, by reference to his own idea of semantic compositionality, which is 'minimal' and is supposedly superior to a so-called truth-conditional idea (i.e. one which portrays a language's structure in a definition of truth for its sentences.) This other style of defence suggests that Horwich's own deflationist attitude may rely on more than the thought that one says everything a philosopher

Pressure can be put upon the minimal theorist, by showing that his conception of truth will be found wanting unless it is supplemented at least with some idea of interpretability. The minimal theorist hopes to say everything that needs to be said about truth by allusion to the Equivalence Schema. So he has said nothing about truth except insofar as the Schema is fully intelligible. When the Schema is used to give a theory of *truth*, it is taken for granted that the 'that *p*' construction within it is in good order. And this might be doubted.[19] But even when there are no doubts about the admissibility of the construction it uses, it must be a good question to ask of someone who advances the Schema what the intended range of its instances is. If one wrote the left hand side of the Schema 'The thinkable that *p* is true', then it would be clear that the 'that p' construction employed in it is just that which is employed when any of 'believes'. 'hopes', 'states' *et cetera* are used. Making out the Schema's intelligibility, or revealing its range, then, would be a matter of showing how 'that p' and related constructions work, as it were, in the large. One would then be drawn to the idea of an interpretive account.[20]

The place to which one turns to answer questions about the minimal theorist's Equivalence Schema, then, is exactly the place where the identity theorist finds the resources to demonstrate her theory's consonance with commonsense. It is not, however, that the allusion to interpretive accounts enables the identity theorist to understand the 'that *p*' construction in terms of something more primitive. Like the minimal theorist, the identity theorist takes the 'that *p*' construction for granted in the statement of a theory of truth. Both theorists assume that

---

could have any reason to say about truth by writing (E) down. Horwich seems sometimes to side with Rorty, see n.21.

    There are other ways of encouraging a generally deflationist attitude than Horwich's. Sometimes it is encouraged by the thought that 'true' having been shown (supposedly) to have no substance, we cannot be 'factualists' about the semantical. (Paul Boghossian, 1990, points out that it is hard to make sense of this except by supposing that we are factualists sometimes; but if we are factualists sometimes, then we cannot actually sustain the wholesale deflationary attitude to truth. Boghossian's thought suggests a difficulty for the move that Horwich makes on my reading-from 'Truth has no substance' to 'Meaning doesn't either'. The move seems to require connections between truth and meaning, which should, it would seem, have been officially renounced.)

[19] Davidson doubts it: see n. 13. Davidson's objections to the minimal theory which were made in his Jacobsen Lectures (University of London, 1995) began from an objection to the use of 'that *p*' in the theorist's Schema.

[20] Such an account, of course, treats utterances of sentences which are not appropriate substitutes for 'p' in Schema (E). (The treatment does not fall within the scope of the definition of truth for a language which it contains, but makes connections with the definition. Consider non-indicatives.) Horwich acknowledges that not all sentences are fit for substitution in (E) at the point at which he glibly dismisses the semantic paradoxes (pp. 41-2).

thinkables are the primary bearers of truth, and give priority to 'It is true that $p$' over 'sentence $s$ is true' (see n. 3). The identity theorist nonetheless sees a particular point in the use of 'true in L' applied to sentences. As we saw, an interpretive account of a particular language L contains a definition of truth in L which speaks to *potential* uses of the sentences of L – to what would be said if one were used by an L-speaker (see §11.3). Within such an account, we find, for instance 'sentence $s$ is true in L iff $p$'; and where this assists in interpretation, the place of '$p$' is taken by a sentence in a theorist's language, fitted for the expression of a thinkable that a speaker would express if, as a speaker of L, she used the sentence s. (The thinkable that $p$ is true if and only if it is a fact that $p$, which of course is exactly the condition for $s$'s being true in L.) An interpretive account enables thinkables which might be expressed by speakers to be matched with thinkables that theorists can express. The idea of such an account would fill the lacuna in the minimal theorist's development of his conception of truth. With the lacuna filled, deflationary impulses go away.[21]

3.3 The identity theorist, unlike the minimal theorist, is willing to look to accounts of interpretation to elicit our conception of truth. But the identity theorist accepts the idea which the minimal theorist's Schema E may be taken to encapsulate. This is the idea that in introducing a word for *truth* into a language not containing such a word, one introduces

---

[21] To subdue the deflationist impulse, it is not enough to acknowledge that definitions of truth for languages play a role in interpretive accounts: the exact role needs to be understood. Rorty promotes a kind of deflationism from the position of a pragmatist, rather than a minimalist. The basis for his dismissive remarks about truth are not the Equivalence Schema, but glosses on truth of the kind that Putnam gives in defending Internal Realism (see n. 14). Speaking of a definition of truth for a language playing the role to which Davidson wishes to put it, Rorty says: 'I should think that an empirical theory which entails T-sentences could as well be called "a theory of complex behaviour" as "a theory of truth"' (1995, p. 286). Well, of course a definition of truth for a language (often enough called 'a theory of truth', which Rorty here expresses a dislike for calling it) is not a theory of truth in the sense in which the correspondence theory, or identity theory, or epistemic theory, or whatever, claim to be. (It is because of the ambiguity in 'theory of truth' that I have here reserved '*definition*' for the Tarskian theories whose role is at issue between identity theorists and minimal theorists: see n.11.) But Rorty's thought that it could 'just as well be called "a theory of complex behaviour"'. does not depend upon any ambiguity in 'theory of truth'. It depends upon Rorty's failure to appreciate that a definition of truth for some speakers' language is [a] only a component in any account of their behaviour, but [b] a component in which 'true' plays a crucial role.

Where Horwich would have us think that the presence of 'true' ensures that the definition participates in nothing with any substance, Rorty, with the focus on 'behaviour', would prevent us from seeing 'true' as playing its norrnal, normative role there (cp. McDowell 1994, p. 150).

nothing which is new (so to speak).[22] For sure, the identity theorist introduces the word 'facts'; but that is only in order to find a way of saying that the facts are the same as what is true, and thus to be quite explicit about her opposition to a correspondence theory of truth. By using the correspondence theorists' own vocabulary of 'fact', the identity theorist sets herself apart from the correspondence theorist's conception of truth. We may ask what the minimal theory is supposed to be set apart from. What might one have wanted to say about truth which it becomes impossible to say when one's conception of truth has been deflated sufficiently to conform to the conception that a minimal theorist recommends?

Well, there are three popular statements of the deflationary idea. First, it is said that truth has no underlying nature. Secondly, it is said that in using the predicate 'true' we cannot get at any more facts than we had access to already. Thirdly, it is said that there is no more to truth than its serving a logical need.

The first claim, that truth has no underlying nature, is Horwich's favourite. But if we try to read it in a distinctively deflationary way, then it may seem to suggest that truth is not a predicate of thinkables with any application independent of other predicates of thinkables. And that is surely wrong.[23]

The second claim, that we cannot get at any more facts than we already have access to, is correct if it makes a point about what the word 'true' enables us to say or to think. (The identity theory shows what is correct about this. When we say or think that $p$ is true, we say or think that it is a fact that $p$ [and it is a fact that $p$ if we are right]; but the fact that $p$ which we may take ourselves now to have access to is the same as

---

[22] Not only is nothing new introduced; nothing is excluded. This point may be used to dispute the title of 'minimal' to the minimal theory. As Gupta 1993 points out, the ideology of the theory is, in one obvious sense, maximal: cp. Frege's remark quoted at n. 7. 23.

[23] Truth's lack of 'an underlying nature' seems sometimes to amount to the impossibility of analyzing or defining it. Such impossibility, of course, is acknowledged by the identity theory: see n. 5.

Horwich sometimes tells us that the predicate 'lacks substance'. But again I find a difficulty. It seems as if *truth* can have any of the features of any property predicated in a true thinkable (cp. nn. 7 & 22). But in that case 'true' is exactly as substantial as, and exactly as insubstantial as, any predicate in the language; which would seem to ensure that we actually cannot make sense of the claim that truth lacks substance. (This peculiarity of 'true' might be used to motivate a treatment in which 'is true' is not a predicate. I do not think that someone with the identity theorist's motivation has to be opposed to such treatment. The kind of deflationism to which it leads need not be assimilated to Horwich's: see n. 16.)

the thinkable that we say or think when we say or think that *p*: there is then no new fact to which we have access.) When a distinctively deflationary understanding of this claim is sought, however, it becomes all too easy to construe it as saying something incorrect. It could be understood as saying that the facts are exhausted by the thinkables to which we already have access. But then it would be suggested that we here now are not entitled to our view that there are facts which no-one here now can actually think. That suggestion is incompatible with the part of commonsense realism whose compatibility with an identity theory I have tried to show.

What, then, about the deflationist's third claim, that there is no more to truth than its serving a certain logical need? This is a claim about the point of having a word which functions as the English word 'true' does. Understood as such, it is correct. But a philosopher who thought to ask 'What is *truth*?' may not be satisfied by being told what sort of device the truth predicate is: she wants to know what sort of distinction it records.

We saw that a grasp of the distinction involved in assessments of thinkables as true or false is required in order to be in a position to make psychological predications of them (§II.3).[24] This is quite compatible with the truth predicate's being the sort of device that the deflationist says it is. And it ensures that there is no new mode of evaluation for thinkables which is brought onto the scene when the word 'true' is introduced. Yet a minimal theorist purports to advance a deflationary claim in using the Equivalence Schema (E) to tell us what sort of device 'true' is. How could he deflate truth further – beyond what is necessary to restore it to the uninflated condition in which the identity theory finds it? It seems that he must either take the distinction involved in assessments of thinkables as true or false to be implicit in some gloss on (E) that he might offer, or he must deny that there is any such distinction. But of course there is a distinction between what is true and

---

[24] There could be debate about whether this suffices to ensure that truth is 'explanatory'. But however that debate might go, truth will not turn out to have the kind of 'causal explanatory role' that e.g. Hartry Field associates with assignments of truth-conditions which count as correspondence truth-conditions. (I touched on some of the issues about truth's role in my 1989.)

what is false. And unless a gloss on (E) has the sort of platitudinous ring that 'fact' carries, it will be bound to spoil the deflationary message.[25]

We have found no positively deflationary and correct thesis about truth for a minimal theorist to advance.

3.4 The conception of truth which the identity theory brings with it allows truth to be a *sui generis* norm, in play where there are rational beings who may go right or wrong in their thought and speech. This is enough to let us shun correspondence theories, and it leaves us well-placed to make out our title to commonsense realism (if that should seem necessary). To the extent that advocates of a minimal theory say distinctively and genuinely deflationary things, they deny us that title; and then they lead us astray.

Answers to philosophers' questions about the relation between language and the world have traditionally taken a form which we now call theories of truth. I have not meant to develop any new theory here. Indeed I do not think that we need any theory of truth save insofar as we may go astray without one. I have promoted the identity theory, because I think that we have to find a position from which to avoid the false dilemmas that the theories currently on offer present us with.[26] I hope that reflection on the identity theory shows that antagonism towards correspondence theories, and indeed towards all theories which purport to analyze truth, is independent of the deflationary attitude. The identity theory provides a perspective from which many other theories will appear indefensible.[27]

---

[25] Wright 1992, Ch. l §1II, argues, in effect, that the deflationist's position is not made out until such a gloss is allowed, but that any such gloss must be inflationary by the deflationist's own lights.

[26] David furnishes a good example of the false dilemma: the reader of David 1994 is invited to accept a correspondence theory of truth on the basis of a demonstration of the untenability of disquotationalism. The dilemmas are sometimes subtler: 'robustness', for instance, is sometimes taken to accrue to truth, or 'factualism' to a discourse which is 'truth apt' as soon as minimalism is denied; and correspondence conceptions may enter with talk of robustness and factualism.

[27] A visit to Monash University in 1991 provided me with some time for thought about truth: for that and more, including a useful discussion of this paper's first ancestor, I thank the Philosophy Department here. Both before and after that visit, I had helpful conversations with Julian Dodd; I can now reciprocate the generous acknowledgement of me in his doctoral thesis. I am grateful both to the members of a discussion group in Oxford which David Charles convenes, and to the Centro de Studie sulla Filosofia Contemporanea for sponsorship of a conference in Genoa in November 1995. My final

REFERENCES

Blackburn, S. 1984: *Spreading the Word*, Oxford: Clarendon Press.

Boghossian, P. 1990: 'The Status of Content', *Philosophical Review* 99, pp. 157- 84.

Brandom, R. 1994: *Making it Explicit*, Cambridge Mass.: Harvard University Press.

Candlish, S. 1995: 'Resurrecting the Identity Theory of Truth', *Bradley Studies*, 1, pp. 116-124.

David, M. 1994: *Correspondence and Disquotation*, Oxford: University Press.

Davidson, D. 1967: 'Truth and Meaning', reprinted as Essay 2 in Davidson 1984.

Davidson, D. 1969: 'True to the Facts', reprinted as Essay 3 in Davidson 1984. Davidson, D. 1984: *Inquiries into Truth and Interpretation*, Oxford: Clarendon Press.

Davidson, D. 1990: 'The Structure and Content of Truth', *Journal of Philosophy* 87, pp. 279-329.

Davies, M. 1981: *Meaning, Quantification, Necessity*, London: Routledge and Kegan Paul.

Dodd, J. 1995: 'McDowell and Identity Theories of Truth', *Analysis* 55, pp. 160- 165.

Dodd, J. & Hornsby, J. 1992: 'The Identity Theory of Truth: Reply to Baldwin', *Mind* 101, pp. 319-322.

Dummett, M. 1958/9: 'Truth', *Proceedings of the Aristotelian Society* 59, pp. 141- 162.

Dummett, M. 1976: 'What is a Theory of Meaning? II', in *Truth and Meaning: Essays in Semantics*, eds. G. Evans & J. McDowell, Oxford: Clarendon Press.

Dummett, M. 1990: 'The Source of the Concept of Truth', in George Boolos (ed.) *Meaning and Method: Essays in Honour of Hilary*

*Putnam*, Cambridge: University Press. [Page references are to the version reprinted in Michael Dummett, The Seas of Language, Oxford: University Press, 1993].

Etchemendy, J. 1988: 'Tarski on Truth and Logical Consequence', *Journal of Symbolic Logic* 53, pp. 51-79.

Field, H. 1986: 'The Deflationary Conception of Truth', in G. MacDonald and C. Wright (eds.) *Fact, Science and Morality: Essays on A. J. Ayer's Language, Truth and Logic*, Oxford: Blackwell.

Frege, G. 1918: 'The Thought'. [Quotations are taken from the translation by A.M. and Marcelle Quinton, as reprinted in *Philosophical Logic*, ed. P.F. Strawson Oxford: Oxford University Press, 1967.]

Grover, D. 1992: *A Prosentential Theory of Truth*, Princeton N.J.: Princeton University Press.

Gupta, A. 1993: 'Minimalism', *Philosophical Perspectives* 7, pp. 359-369.

Hornsby, J. 1989: 'Semantic Innocence and Psychological Understanding', *Philosophical Perspectives*, 3, pp. 549-574. Horwich, P. 1990: Truth, Oxford: Blackwell.

Jackson, F., Oppy G. & Smith, M. 1994: 'Minimalism and Truth Aptness', *Mind* 103, pp. 287-302.

McDowell, J. 1994: *Mind and World* (Cambridge Mass.: Harvard University Press).

McDowell, J. 1976: 'Bivalence and Verificationism', in Gareth Evans and John McDowell (eds.) *Truth and Meaning: Essays in Semantics*, Oxford: Clarendon Press.

Neale, S. 1995: 'The Philosophical Significance of Godel's Slingshot', *Mind* 104, pp. 761-825.

Putnam, H. 1981: *Reason Truth and History*, Cambridge: University Press.

Putnam, H. 1990: 'A Defense of Internal Realism', in *Realism with a Human Face*, Cambridge Mass.: Harvard University Press.

Quine, W.V. 1960: *Word and Object*, Cambridge, Mass.: The M.I.T. Press.

Quine, W.V. 1970: *Philosophy of Logic*, Englewood Cliffs: Prentice-Hall.

Rorty, R. 1995: 'Is Truth a Goal of Enquiry? Davidson vs Wright', *Philosophical Quarterly* 45, pp. 281-300.

Rumfitt, I. 1993: 'Content and Context: The Paratactic Theory Revisited and Revised', *Mind* 102, pp. 429-454.

Russell, B. 1940: *An Inquiry into Meaning and Truth*, London: Allen and Unwin.

Wright, C. 1992: *Truth and Objectivity*, Cambridge Mass.: Harvard University Press.

*Introduction to and Commentary on
Jennifer Hornsby's
'Truth: The Identity Theory'"*

GILA SHER

# INTRODUCTION TO AND COMMENTARY ON JENNIFER HORNSBY'S 'TRUTH: THE IDENTITY THEORY'

## GILA SHER

JENNIFER Hornsby's 1997 paper, 'Truth: The Identity Theory', has been highly influential in making the identity theory of truth a viable option in contemporary philosophy. In this short introduction and commentary I will limit my attention to what distinguishes her theory and its methodology from the correspondence theory and the 'substantivist' methodology, and I will focus on issues that have not been widely discussed in earlier commentaries yet are central to the current debate on truth.[1] By 'the identity theory' I will mean 'Hornsby's identity theory'.

I. *Introduction.* The identity theory says that a truth-bearer is true if and only if it is a fact or is the *same as* some (appropriate) fact. For Hornsby the justification or *raison d'être* of the identity theory is mainly negative: 'the identity theory arises out of rejection of a correspondence theory' (p. 4); '[t]he interest of the theory derives from what it [is] opposed to philosophically' (p. 3); 'the identity theory is worth considering to the extent to which correspondence theories are worth avoiding' (p. 6). The key issue is the relation between truth-bearers and reality. The correspondence theory says that there is a 'gap' between truth-bearers (thoughts) and something *external* to them which explains their truth/falsehood. The identity theory says there is no such gap.

Hornsby could have tried to justify her theory by appealing to its avoidance of criticisms like the 'slingshot', directed at traditional correspondence theories. But although she says in a footnote that the slingshot criticism does not apply to her theory, she does not pursue this line of justification. She focuses on a more central issue to the correspondence theory, namely, whether truth requires a 'gap' between truth-bearers and reality, and she motivates the identity theory by objections, which she attributes to Frege (1918), Quine (1960) and McDowell (1994), to the 'gap' view.

Frege argued that truth does not come in degrees or 'respects'; truth requires a *perfect* correspondence between thought and reality, and perfect correspondence is *coincidence.* Hornsby interprets him as saying that 'there cannot be an ontological gap between thought ("an idea") and the world ("something real")' (Hornsby 1997, p. 6). Next, she

---

[1] For other contemporary identity theorists, commentators on Hornsby, and Hornsby's responses see, e.g., Candlish (1999), Hornsby (1999), Dodd (1999, 2000), David (2001) and Engel (2001).

traces the 'no gap' view to Quine's attack on the *myth of the museum*: '[f]rom the identity theorist's point of view the correspondents of the correspondence theories of truth play the same role as the exhibits in the museum of the mind'. Most importantly for the contemporary debate, Hornsby traces the 'no gap' view to McDowell:

> [T]here is no ontological gap between the sort of thing one can ... think, and the sort of thing that can be the case. When one thinks truly, what one thinks *is* the case .... [T]here is no gap between thought, as such, and the world. [McDowell 1994, p. 27][2]

The identity theory is concerned with one aspect of the 'gap' question: whether there is a gap between *truth-bearers* and *facts*. One distinctive characteristic of the theory is its conception of truth-bearers as *thinkables*, a notion borrowed from McDowell. Thinkables, as Hornsby understands them, are neither mental nor worldly entities. They are *contents* of thoughts - entities that may be thought of as located in a Fregean realm of sense. 'The identity theory is encapsulated in the simple statement that true thinkables are the same as facts' (p. 2). By identifying facts with true thinkables (= true truth-bearers), the world plays no role in what the identity theory has to say about truth: 'there is nothing external to thought [truth-bearers, thinkables] ... in terms of which truth can be understood' (p. 17).

As a theory, the identity theory is anti-substantivist. Among the things it is designed *not to do* are give a definition of truth, analyze truth, provide an explanation of truth, look for the sources of truth, say anything which is metaphysically contentious, examine what people are doing when they investigate reality, and so on. 'Truth', according to the identity theory, is not a notion of substance that can be used in explaining other things. By introducing this notion into our language we do not add anything new to it. There is no more to truth than playing a certain technical (logical) role. Truth does not provide a new mode of evaluating truth-bearers. From the point of the view of the identity theory even Horwich's minimalist theory tries to do too much: 'to the extent that the minimal theorist wants to convey a deflationary message about truth, which is not already conveyed in the identity theorist's opposition to correspondence, the message has to be resisted' (Hornsby 1997, p. 16). This does not mean that the identity theory is vacuous: it takes a stand on what the bearers of truth are and, according to Hornsby, it also has resources for saying what people are doing when they are using language and for offering an interpretive account of truth. But it eschews everything concerning truth that goes beyond speakers and truth-bearers.

---

[2] Others who connect the identity theory of truth with McDowell (1994) include Dodd (1995), Suhm, Wgemann & Wessels (1999), Engel (2001), McDowell (2005) and Fish & Mcdonald (2007).

Hornsby does not justify her anti-substantivist approach to truth beyond pointing to Frege's (1918) claim that truth is indefinable (which is itself justified on the grounds that due to the basicness of truth, a definition of this notion is bound to involve circularity or infinite regress). Possibly, she also obliquely appeals to McDowell's reasons for recommending quietism in philosophy; possibly, her anti-substantivism is rooted in her overall approach to philosophy, which emphasizes commonsense over theory. But there is no explicit justification of anti-substantivism in Hornsby's paper.

II. *Commentary.* By appealing to McDowell (1994) in motivating her theory of truth, Hornsby (along with Dodd 1995) made an original and important contribution to the contemporary debate on truth. The reason is that the basic problematics of the human cognitive situation that McDowell draws attention to is both central to truth and largely neglected in the current debate. Hornsby's weakness is that she limits her attention to one element of this problematics, the 'no gap' element, losing sight of the equally important 'gap' element, hence of the problematics itself. Let me explain:

McDowell calls for following Kant in recognizing the problematics of the human cognitive situation. This situation, as McDowell delineates it, is characterized by several polarities: mind and world, friction and freedom, concept and object. For true cognition to occur, '[h]uman minds must somehow be able to latch on to the inhuman structure of reality' (McDowell 1994, p. 77). Now, such a latching on requires two elements, a human *mind* and a *world* independent of it, hence *gap*. Latching on to the world is bridging or closing the gap (hence *no gap*). Both are essential for true cognition. In true cognition mind must be *constrained* by the external world (*gap*), but in a way that makes it possible for it to use the constraint to generate true cognition as distinguished from false one. McDowell's proposal is that the world, or that part of the world that constrains the mind in this constructive manner, is *thinkable*. That is, the constraint involved in true cognition, which is rational or conceptual, is itself *rational*, that is, involves, or is mediated by, *concepts*. Concepts, in turn, require *freedom*, hence true cognition has a substantial element of freedom. More specifically, McDowell proposes that concepts play a central role in the mind's latching on to the world *all the way*. That is, there is no gap between the use of concepts in latching on to the world and the use of pure perception. The kind of perception that is relevant for cognition is already imbued with concepts. This is the more specific meaning of the '*no gap*' element in McDowell's proposal. But McDowell is adamant that we must not neglect the crucial role of *external* constraint in this process (hence *gap*). Constraint by the external world (*gap*), he repeats time and again, is crucial for true cognition. Without it, what we regard as true cognition would be a mere 'spectre of a frictionless spinning in a void' (p. 18).

How is this relevant to the theory of truth? It is relevant in two directions: (a) The question of truth is central to the cognitive problem since the difference between failing and succeeding in latching on to the world is the difference between obtaining *false* and *true* cognitions. (b) The question of cognition is central to truth, since one of the central roles of truth is to distinguish between failed and successful cognitions of the world (latching-ons to the world), or set a standard for successful cognitions. Accordingly, a theory of truth cannot neglect external constraint by the world. But this crucial aspect of McDowell's conception is entirely missing from the identity theory of truth. External constraint plays no role in this theory.

Another way to arrive at this issue is to observe that if a theory of truth puts McDowell's cognitive problematics at its center, it faces a non-trivial challenge: How to account for the truth of human thoughts given the enormous complexity of the human cognitive situation. (How to account for the fact that true thoughts are products of the mind yet their truth is determined by the external world;[3] how to deal with the polarities involved in the cognitive situation, which pull us in seemingly opposite directions.) The problem is further magnified since, as McDowell rightly demands, we must avoid the traps of idealism on the one hand and a Kantian thing-in-itself / empiricist 'brute Given' on the other. But the identity theory cannot meet this challenge. First, identity, as a gap-denying relation, can play at most a limited role in an account involving gaps and non-symmetric relations (like 'latching on' to the world). Second, given the complexity of the cognitive situation involved in truth, it is quite unlikely that a theory as simple as the identity theory has sufficient resources to account for it. Third, given the importance of the McDowellian problematics for understanding true cognition (truth), it requires a substantive treatment, something the identity theory is not equipped to provide.[4]

Hornsby is quite aware of the fact that truth involves an independent reality (hence gap); indeed, she regards it as a virtue of her theory that it is compatible with commonsense realism. She is also aware that there is an important asymmetry in truth (which cannot be captured by identity), and that this asymmetry concerns an important aspect of truth, namely, that it is reality which determines the truth of our thoughts, rather than our thoughts which determine reality (Hornsby 2005, §4). But her theory does not account for any of these things, nor do any of these things play a substantial role in her theory. Why? I gather this is because she realizes that accounting for these characteristics of truth would require a *substantive* theory or truth and, moreover, a *correspondence*

---

[3] To see that, and how, McDowell views the world as determining the truth/falsehood of a particular truth-bearer see his example in (2005, p. 85).

[4] In referring to McDowell's problematics, I focus primarily on the early lectures in his book. Whether any of my critical points applies to any of McDowell's views elsewhere in his corpus I will not be able to discuss here.

theory of truth (or something like it), but (i) she is anti-substantivist with respect to truth, and (ii) she sees no hope for a correspondence-like theory of truth.

The key to understanding the problem with Hornsby's attitude (namely (i) and (ii)) is provided by Hornsby herself. She concludes her paper by saying: 'I have promoted the identity theory, because I think that we have to find a position from which to avoid the false dilemmas that the theories currently on offer present us with' (p. 22). To see our way into a theory that meets the McDowellian challenge we must free ourselves from the false misconceptions and dilemmas that are commonly associated with theories of this kind.

Consider correspondence. From the point of view of the basic problematics of human cognition correspondence is a relation that seriously involves the human mind on the one hand and the world on the other, a relation that explains how the world constrains the mind in a way that gives rise to true cognition. Now, from the point of view of these concerns, whether the basic entities involved in this relation are *facts* or entities of other kinds (say, objects and properties) is a secondary, and *open*, question. But from the point of view of the *traditional* debate on correspondence the question of facts is the main question. Viewed traditionally, correspondence requires *wholes* of one kind to stand in some fixed relation to *wholes* of another kind, and the question is which *wholes* these are? But what Hornsby does not seem to see is that from the point of view of the McDowellian problematics the question of *wholes* or *facts* is merely scholastic. What we are interested in is *how* the mind latches on to the world to achieve true cognition. That is, our investigation is directed at the *routes* the mind can, does, and should, take into the world in order to attain true cognition. And this mean that it is an open question whether this route leads from whole sentences (whole thoughts) to whole pieces of the world (facts) or whether it proceeds by connecting language (thought) to objectual elements that are essentially *different* from facts. But the objections to correspondence mentioned by Hornsby do not apply to a (non-traditional) correspondence theory that does not require facts, a theory that *investigates* the correspondence relation and its relata rather than takes them as given.

In a way, Hornsby cannot consider a correspondence investigation of this kind, since such an investigation would in all likelihood be substantive, but Hornsby objects to a substantive study of truth. Hornsby might argue that nonsubstantive theories have some advantages over substantive theories. For example, a nonsubstantive theory is more "safe" than a substantive theory in the sense that it says very little about truth and as such it is unlikely to conflict with most of the true things that can be said about truth, hence is unlikely to be found incorrect. But is this really an advantage? McDowell is known for emphasizing the importance of not confusing *justification* with *exculpation*. Saying so little about truth that one cannot be wrong about anything substantive

concerning it grants one *immunity to blame*. But what a good theory needs is justification. A theory has to say enough so that on the one hand it is open to critical examination and on the other hand it teaches us something new about its subject matter – here, the structure of truth, given the complexities of the human cognitive situation.

Perhaps, however, a substantive theory of truth is impossible. Hornsby, as we have seen, points to Frege's claim that such a theory cannot avoid circularity. Is this a decisive objection? I think it is not, and the key to understanding why it is not is, following Hornsby's own counsel of freeing ourselves from false dilemmas and preconceptions. The view that any measure of circularity is fatal to a theory is tied to a specific methodology, *foundationalism*. The first step in overcoming Frege's objection is recognizing the availability of alternative philosophical methodologies, like *holism*, which sanction some measure and forms of circularity (and is, indeed, endorsed by McDowell). Holism, however, is traditionally assumed to be a coherentist methodology, incompatible with correspondence. The second step is to realize the falsity of this assumption. Holism licenses non-linear theorizing in principle, and this license can be used in a variety of theories, including theories directed at, and grounded in, reality. Holism focuses not on the order in which such theories are constructed but on the extent to which they contribute to our knowledge of a given subject-matter. It allows shifts in position within our system of knowledge, conceived as a Neurath boat, including shifts involving temporary and partial circularity (circularity that can eventually be eliminated, circularity that involves only parts of the theory, and circularity whose extent can be reduced in later stages of developing the theory). As such, it is especially suited to the study of subject-matters like truth. The viability of a substantive theory of truth using the holistic method is not undermined by Frege who had never considered this possibility.

There is much more to say about the issues discussed in this commentary, but my space is limited.[5] To clarify my criticism of Hornsby's paper, let me briefly respond to a question raised by Guy Longworth. Longworth suggested that there are two ways to understand my criticism: (1) The claim is that Hornsby does not provide a substantive account of the gap and the asymmetry involved in truth. (2) The claim is that her theory does not have the resources to provide such an account. My view is that if it is inherent in Hornsby's approach that the Identity Theory is anti-substantivist (as I believe it is), then the criticism is (2). If, on the other hand, Hornsby's conception of the Identity theory is compatible with a substantivist approach to truth, then the criticism is (1). In the latter case, the challenge I put to Hornsby is to

---

[5] For more on (non-coherentist) holism, the substantivist approach to truth, and 'routes' of correspondence see Sher (1999, 2004, 2010, 2012, 2013a,b, and Forthcoming).

provide a substantive account of the gap and asymmetry involved in truth, one that is significantly an "identity" account.

The questions of substantiveness, correspondence, the relation between truth and cognition, and the focus of theories of truth are important questions that many of the seminal papers published in the *Proceedings of the Aristotelian Society* bear upon. The identity theory of truth falls on one side of these issues; I support the other. I hope that the problematics I focused on in these comments will stimulate a fruitful debate on these issues.

REFERENCES

Candlish, S. 1999. 'Identifying the Identity Theory of Truth'. *Proceedings of the Aristotelian Society*, New Series. 99: 233-40.

David, M. 2001. 'Truth as Identity and Truth as Correspondence'. *The Nature of Truth: Classic & Contemporary Perspectives.* Cambridge, MA: MIT. Pp. 683-704.

Dodd, J. 1995. 'McDowell & Identity Theories of Truth'. *Analysis* 55: 160-5.

----. 1999. 'Hornsby on the Identity Theory of Truth'. *Proceedings of the Aristotelian Society*, New Series. 99: 225-32.

----. 2000. *The Identity Theory of Truth.* Houndmills: Macmillan.

Engel, P. 2001. 'The False Modesty of the Identity Theory of Truth'. *International Journal of Philosophical Studies* 9: 441-58.

Fish, W. & C. Macdonald. 2007. *Analysis* 67: 36-41.

Frege, G. 1918. 'The Thought'. *Logical Investigations.* Ed. P.T. Geach. Oxford: Basil Blackwell, 1977. Pp. 1-30.

Hornsby, L. 1997. 'Truth: The Identity Theory'. *Proceedings of the Aristotelian Society*, New Series. 97: 1-24.

----. 1999. 'The Facts in Question: A Response to Dodd and to Candlish'. *Proceedings of the Aristotelian Society*, New Series. 99: 241-5.

----. 2005. 'Truth without Truthmaking Entities'. *Truthmakers: The Contemporary Debate.* Eds. H. Beebee & J. Dodd. Oxford, Pp. 33-48.

McDowell, J. 1994. *Mind & World.* Cambridge, MA: Harvard.

----. 2005. 'The True Modesty of an Identity Conception of Truth: A Note in Response to Pascal Engel (2001)'. *International Journal of Philosophical Studies* 13: 83-8.

Quine, V.W. 1960. *Word & Object.* Cambridge, MA: MIT.

Sher, G. 1999. 'On the Possibility of a Substantive Theory of Truth'. *Synthese* 117: 133-172.

----. 2004. 'In Search of a Substantive Theory of Truth. *Journal of Philosophy* 101: 5-36.

----. 2010. 'Epistemic Friction: Reflections on Knowledge, Truth, and Logic'. *Erkenntnis* 72: 151-76

----. 2012. 'Truth & Knowledge in Logic & Mathematics'. *The Logica Yearbook 2011.* Eds. M. Peliš & V. Punčochář. London: College Publications, King's College, Pp. 289-304.

----. 2013a. 'Forms of Correspondence: The Intricate Route from Thought to Reality'. *Truth & Pluralism: Current Debates.* Eds. N.J.L.L. Pedersen & C.D. Wright. Oxford: 157-79.

----. 2013b. 'The Foundational Problem of Logic'. *The Bulletin of Symbolic Logic* 19: 145-98.

----. Forthcoming. 'Truth as Composite Correspondence'. *Unifying the Philosophy of Truth.* Eds. T. Achourioti, H. Galinon, H. Fujimoto, and J. Martinez-Fernández. Dordrecht: Springer.

Suhm, C., P. Wagemann & F. Wessels. 1999. 'Ontological Troubles with Facts & Objects in McDowell's *Mind & World* '. *John McDowell: Reason & Nature.* Lecture & Colloquium. Ed. M. Willaschek. Münster: LIT - Verlag. Pp. 27-33.