BERTRAND RUSSELL

Logical and Philosophical Papers 1909–13

Edited by
John G. Slater
with the assistance of
Bernd Frohmann

Routledge
Taylor & Francis Group
LONDON AND NEW YORK

12

On the Nature of Truth and Falsehood [1910]

THIS ESSAY WAS first published in *Philosophical Essays* (1910), where it had the distinction of being the only complete paper which had not previously been published. There is a suggestion, in a letter from Hugh Elliot, a relative of Russell's on his mother's side, to Russell, dated 21 March 1909, that Russell wrote this essay in order to interest a publisher in bringing out a collection of his philosophical essays:

> I approached Longmans with reference to your collection of essays; and from their reply, enclosed herewith, you will see that they are ready to publish them. Perhaps you would now negotiate with them directly. I told them you were adding another essay, not hitherto published, so they know all about it. In giving them a description of the essays, I made no special mention of that which you published in the Journal of the Aristotelian Society. It did not seem to me at all too hard to understand. I am looking forward very much to reading the volume when it appears.

The article which had appeared in the *Proceedings of the Aristotelian Society* was "On the Nature of Truth" (1907); the first two parts of it were reprinted in *Philosophical Essays* under the title, "The Monistic Theory of Truth". Russell had become dissatisfied with Part III and did not wish to see it reprinted. As a substitute for it he wrote the present paper.

In addition to the printed version, there is the manuscript (RA 220.011190), which has been selected as copy-text. The two versions have been collated; the results are reported in the Textual Notes.

DOI:10.4324/9781003557395-15

Which question: what truth means, not what things are true.

THE QUESTION "what is Truth?" is one which may be understood in several different ways, and before beginning our search for an answer, it will be well to be quite clear as to the sense in which we are asking the question. We may mean to ask what things are true: is science true? is revealed religion true? and so on. But before we can answer such questions as these, we ought to be able to say what these questions *mean*: what is it, exactly, that we are asking when we say "is science true?" It is this preliminary question that I wish to discuss. The question whether this or that is true is to be settled, if at all, by considerations concerning this or that, not by general considerations as to what "truth" means; but those who ask the question presumably have in their minds already some idea as to what "truth" means, otherwise the question and its answer could have no definite meaning to them.

When, however, we have agreed that the question we are concerned with is "what does 'truth' mean?", we have by no means come to an end of possible ambiguities. There is the question: "How is the word 'truth' properly used?" This is a question for the dictionary, not for philosophy. Moreover the word has some perfectly proper uses which are obviously irrelevant to our inquiry: a "true" man, a "true" poet, are "true" in a different sense from that with which we are concerned. Again there is the question: "What do people usually have in mind when they use the word 'truth'?" This question comes nearer to the question we have to ask, but is still different from it. The question what idea people have when they use a word is a question of psychology; moreover there is very little in common between the ideas which two different people in fact attach to the same word, though there would often be more agreement as to the ideas which they would consider it proper to attach to the word.

The question we have to discuss may be explained by pointing out that, in the case of such a word as "truth", we all feel that some fundamental concept, of great philosophical importance, is involved, though it is difficult to be clear as to what this concept is. What we wish to do is to detach this concept from the mass of irrelevancies in which, when we use it, it is normally embedded, and to bring clearly before the mind the abstract opposition upon which our distinction of true and false depends. The process to be gone through is essentially one of analysis: we have various complex and more or less confused beliefs about the true and the false, and we have to reduce these to forms which are simple and clear, without causing any avoidable conflict between our initial complex and confused beliefs and our final simple and clear assertions. These final assertions are to be tested partly by their intrinsic evidence, partly by their power of accounting for the "data"; and the "data", in such a problem, are the complex and confused beliefs with which we start. These beliefs must

necessarily suffer a change in becoming clear, but the change should not be greater than is warranted by their initial confusion.

Truth-bearers: beliefs and judgments, and the sun example.

Although the question what things are true rather than false does not form part of our inquiry, yet it will be useful to consider for a moment the nature of the things to which we attribute either truth or falsehood. Broadly speaking, the things that are true or false, in the sense with which we are concerned, are statements, and beliefs or judgments.[1] When, for example, we see the sun shining, the sun itself is not "true", but the judgment "the sun is shining" is true. The truth or falsehood of statements can be defined in terms of the truth or falsehood of beliefs: A statement is true when a person who believes it believes truly, and false when a person who believes it believes falsely. Thus in considering the nature of truth, we may confine ourselves to the truth of beliefs, since the truth of statements is a notion derived from that of beliefs. The question we have to discuss is, therefore: What is the difference between a true belief and a false belief? By this I mean, what is the difference which actually *constitutes* the truth or falsehood of a belief? I am not asking for what is called a *criterion* of truth, i.e. for some quality, other than truth, which belongs to whatever is true and to nothing else. This distinction between the *nature* of truth and a *criterion* of truth is important, and has not always been sufficiently emphasized by philosophers. A criterion is a sort of trade-mark, i.e. some comparatively obvious characteristic which is a guarantee of genuineness. "None genuine without the label": thus the label is what assures us that such-and-such a firm made the article. But when we say that such-and-such a firm made the article, we do not *mean* that the article has the right label; thus there is a difference between meaning and criterion. Indeed it is just this difference which makes a criterion useful. Now I do not believe that truth has, universally, any such trade-mark: I do not believe that there is any one label by which we can always know that a judgment is true rather than false. But this is not the question which I wish to discuss: I wish to discuss what truth and falsehood actually are, not what extraneous marks they have by which we can recognize them.

Nature against criterion: the trade-mark, and no universal label.

The first point upon which it is important to be clear is the relation of truth and falsehood to the mind. If we were right in saying that the things that are true or false are always judgments, then it is plain that there can be no truth or falsehood unless there are minds to judge. Nevertheless it is plain, also, that the truth or falsehood of a given judgment depends in no way upon the person judging, but solely upon the facts about which he judges. If I judge that Charles I died in his bed, I judge falsely, not

1 I shall use the words "belief" and "judgment" as synonyms.

The datum: truth needs minds, yet rests on the facts.

because of anything to do with me, but because in fact he did not die in his bed. Similarly, if I judge that he died on the scaffold, I judge truly, because of an event which in fact occurred 260 years ago. Thus the truth or falsehood of a judgment always has an objective ground, and it is natural to ask whether there are not objective truths and falsehoods which are the objects, respectively, of true and false judgements. As regards truths, this view is highly plausible. But as regards falsehoods, it is the very reverse of plausible; yet, as we shall see, it is hard to maintain it with regard to truths without being forced to maintain it also as regards falsehoods.

Cognitive acts have objects, and judging: one object or several?

In all cognitive acts, such as believing, doubting, disbelieving, apprehending, perceiving, imagining, the mind has objects other than itself to which it stands in some one of these various relations. In such a case as perception, this is sufficiently obvious: the thing perceived is necessarily something different from the act of perceiving it, and the perceiving is a relation between the person perceiving and the thing perceived. The same thing holds, though less obviously, with regard to imagination. If I imagine, say, a certain colour, the colour is an object before my mind just as truly as if I perceived the colour, though the relation to my mind is different from what it would be if I perceived the colour, and does not lead me to suppose that the colour exists in the place where I imagine it. Judgments also consist of relations of the mind to objects. But here a distinction has to be made between two different theories as to the relation which constitutes judgment. If I judge (say) that Charles I died on the scaffold, is that a relation between me and a single "fact", namely Charles I's death on the scaffold, or "that Charles I died on the scaffold", or is it a relation between me and Charles I and dying and the scaffold? We shall find that the possibility of false judgments compels us to adopt the latter view. But let us first examine the view that a judgment has a single object.

Objectives, after Meinong: the true case looks fine.

If every judgment, whether true or false, consists in a certain relation, called "judging" or "believing", to a single object, which is what we judge or believe, then the distinction of true and false as applied to judgments is derivative from the distinction of true and false as applied to the objects of judgments. Assuming that there are such objects, let us, following Meinong, give them the name "Objectives". Then every judgment has an Objective, and true judgments have true Objectives, while false judgments have false Objectives. Thus the question of the meaning of truth and falsehood will have to be considered first with regard to Objectives, and we shall have to find some way of dividing Objectives into those that are true and those that are false. In this, however, there is great difficulty. So long as we only consider true judgements, the view that they have Objectives is plausible: the actual event which we describe as "Charles I's death on the scaffold" may be regarded as the Objective of the judgment "Charles I died on the scaffold". But what is the Objective of the judgment

No such event as the death in bed, so the Objective is a 'that'.

"Charles I died in his bed"? There was no event such as "Charles I's death in his bed". To say that there ever was such a thing as "Charles I's death in his bed" is merely another way of saying that Charles I died in his bed. Thus if there is an Objective, it must be something other than "Charles I's death in his bed". We may take it to be "that Charles I died in his bed". We shall then have to say the same of true judgments: the Objective of "Charles I died on the scaffold" will be "that Charles I died on the scaffold".

First objection: 'that so-and-so' doesn't name an object.

To this view there are, however, two objections. The first is that it is difficult to believe that there are such objects as "that Charles I died in his bed" or even "that Charles I died on the scaffold". It seems evident that the phrase "that so-and-so" has no complete meaning by itself, which would enable it to denote a definite object as (e.g.) the word "Socrates" does. We feel that the phrase "that so-and-so" is essentially incomplete, and only acquires full significance when words are added so as to express a judgment, e.g. "I believe that so-and-so", "I deny that so-and-so", "I hope that so-and-so". Thus if we can avoid regarding "that so-and-so" as an independent entity, we shall escape a paradox. This argument is not decisive, but it must be allowed a certain weight.

Second objection: falsehoods in a world without minds.

The second objection is more fatal, and more germane to the consideration of truth and falsehood. If we allow that all judgments have Objectives, we shall have to allow that there are Objectives which are false. Thus there will be in the world entities, not dependent upon the existence of judgments, which can be described as objective falsehoods. This is in itself almost incredible: we feel that there could be no falsehood if there were no minds to make mistakes. But it has the further drawback that it leaves the difference between truth and falsehood quite inexplicable. We feel that when we judge truly some entity "corresponding" in some way to our judgment is to be found outside our judgment, while when we judge falsely there is no such "corresponding" entity. It is true we cannot take as this entity simply the grammatical subject of our judgment: if we judge, e.g., "Homer did not exist", it is obvious that Homer is not the entity which is to be found if our judgment is true but not if it is false. Nevertheless it is difficult to abandon the view that, in some way, the truth or falsehood of a judgment depends upon the presence or absence of a "corresponding" entity of some sort. And if we do abandon this view, and adhere to the opinion that there are both true and false Objectives, we shall be compelled to regard it as an ultimate and not further explicable fact that Objectives are of two sorts, the true and the false. This view, though not logically impossible, is unsatisfactory, and we shall do better, if we can, to find some view which leaves the difference between truth and falsehood less of a mystery.

The further drawback: correspondence lost, and the Homer case.

It might be thought that we could say simply that true judgments have

Two ways round fail: no relation to nothing, no intrinsic mark.

Objectives while false ones do not. With a new definition of Objectives, this view might become tenable, but it is not tenable so long as we hold to the view that judgment actually is a relation of the mind to an Objective. For this view compels us, since there certainly are false judgments, and a relation cannot be a relation to nothing, to admit that false judgments as well as true ones have Objectives. We must, therefore, abandon the view that judgments consist in a relation to a single object. We cannot maintain this view with regard to true judgments while rejecting it with regard to false ones, for that would make an intrinsic difference between true and false judgments, and enable us (what is obviously impossible) to discover the truth or falsehood of a judgment merely by examining the intrinsic nature of the judgment. Thus we must turn to the theory that *no* judgment consists in a relation to a single object.

The way out: several objects, as real for the false judgment.

The difficulty of the view we have been hitherto considering was that it compelled us either to admit objective falsehoods, or to admit that when we judge falsely there is nothing that we are judging. The way out of the difficulty consists in maintaining that, whether we judge truly or whether we judge falsely, there is no one thing that we are judging. When we judge that Charles I died on the scaffold, we have before us, not one object, but several objects, namely Charles I and dying and the scaffold. Similarly when we judge that Charles I died in his bed, we have before us the objects Charles I, dying, and his bed. These objects are not fictions: they are just as good as the objects of the true judgment. We therefore escape the necessity of admitting objective falsehoods, or of admitting that in judging falsely we have nothing before the mind. Thus in this view judgment is a relation of the mind to several other terms: when these other terms have inter se a "corresponding" relation, the judgment is true; when not, it is false. This view, which I believe to be the correct one, must now be further expanded and explained.

Not three relations but one relation of more than two terms.

In saying that judgment is a relation of the mind to several things, e.g. to Charles I and the scaffold and dying, I do not mean that the mind has a certain relation to Charles I and also has this relation to the scaffold and also has it to dying. I do not, however, wish to deny that, when we are judging, we have *a* relation to each of the constituents of our judgment separately, for it would seem that we must be in some way conscious of these constituents, so that during any judgment, we must have, to each constituent of the judgment, that relation which we may call "being conscious of it". This is a very important fact, but it does not give the essence of judgment. Nothing that concerns Charles I and dying and the scaffold separately and severally will give the judgment "Charles I died on the scaffold." In order to obtain this judgment, we must have one single unity of the mind and Charles I and dying and the scaffold, i.e. we must have, not several instances of a relation between two terms, but one instance of

a relation between more than two terms. Such relations, though familiar to mathematicians, have been unduly ignored by philosophers. Since they appear to me to give the key to many puzzles about truth, I shall make a short digression to show that they are common and ought to be familiar.

A loved B in May, and the date as a term of the relation.

One of the commonest ways in which relations between more than two terms occur is in propositions about what happened at some particular time. Take such a proposition as "*A* loved *B* in May and hated him in June", and let us suppose this to be true. Then we cannot say that, apart from dates, *A* has to *B* either the relation of loving or that of hating. This necessity for a date does not arise with *all* ordinary relationships; for example if *A* is the brother of *B*, no date is required: the relationship holds always or never, or (more strictly) holds or does not hold without regard to time. But love and hate are "time's fool": they are not relations which hold without regard to date. "*A* loved *B* in May" is a relation, not between *A* and *B* simply, but between *A* and *B* and May.[2] This relation between *A* and *B* and May cannot be analyzed into relations between *A* and *B*, *A* and May, *B* and May: it is a single unity. It is partly the failure to perceive that the date is one of the terms in such relations which has caused such difficulty in the philosophy of time and change.

Jealousy is triangular: the Othello case.

As another illustration, take the relation of jealousy. Time comes in here exactly as it did with love and hate, but we will for the moment ignore time, because the point to be noticed about jealousy is that it involves three people. The simplest possible proposition asserting jealousy is such as "*A* is jealous of *B*'s love for *C*", or "*A* is jealous of *B* on account of *C*". It might be thought that "*B*'s love for *C*" was one term, and *A* the other term. But this interpretation will not apply to cases of mistaken jealousy: if *A* is Othello, there is no such thing as "*B*'s love for *C*". Thus this interpretation is impossible, and we are compelled to regard jealousy as a relation of three persons, i.e. as having for its unit a relation which is what we may call "triangular". If we further take into account the necessity for a date, the relation becomes "quadrangular", i.e. the simplest possible proposition involving the relation will be one which concerns four terms, namely three people and a date.

The definition of a multiple relation.

We will give the name "multiple relations" to such as require more than two terms. Thus a relation is "multiple" if the simplest propositions in which it occurs are propositions involving more than two terms (not counting the relation). From what has been said, it is obvious that multiple relations are common, and that many matters cannot be understood

2 I do not want to assume any theory as to the nature of time: "May" can be interpreted as the reader likes. The statement in the text may then have to be made a little more complicated, but the necessity for a relation of more than two terms will remain.

without their help. Relations which have only two terms we shall call "dual relations".

The theory stated on A loves B, and truth as a first approximation.

The theory of judgment which I am advocating is, that judgment is not a dual relation of the mind to a single Objective, but a multiple relation of the mind to the various other terms with which the judgment is concerned. Thus if I judge that *A* loves *B*, that is not a relation of me to "*A*'s love for *B*", but a relation between me and *A* and love and *B*. If it were a relation of me to "*A*'s love for *B*", it would be impossible unless there were such a thing as "*A*'s love for *B*", i.e. unless *A* loved *B*, i.e. unless the judgment were true; but in fact false judgments are possible. When the judgment is taken as a relation between me and *A* and love and *B*, the mere fact that the judgment occurs does not involve any relation between its objects *A* and love and *B*; thus the possibility of false judgments is fully allowed for. When the judgment is true, *A* loves *B*; thus *in this case*, there is a relation between the objects of the judgment. We may therefore state the difference between truth and falsehood as follows: Every judgment is a relation of a mind to several objects, one of which is a relation; the judgment is *true* when the relation which is one of the objects relates the other objects, otherwise it is false. Thus in the above illustration, love, which is a relation, is one of the objects of the judgment, and the judgment is true if love relates *A* and *B*. The above statement requires certain additions which will be made later; for the present, it is to be taken as a first approximation.

Why perception can't err: dreams as errors of judgment.

One of the merits of the above theory is that it explains the difference between judgment and perception, and the reason why perception is not liable to error as judgment is. When we were considering the theory that judgment is a dual relation of the mind to a single Objective, we found that so far as true judgments were concerned this theory worked admirably, but that it would not account for false judgments. Now this difficulty will not apply against a corresponding theory of perception. It is true that there are cases where perception *appears* to be at fault, such as dreams and hallucinations. But I believe that in all these cases the perception itself is correct, and what is wrong is a judgment based upon the perception. It would take us too far from our subject to develop this theme, which requires a discussion of the relation between sense-data (i.e. the things we immediately perceive) and what we may call physical reality, i.e. what is there independently of us and our perceptions. Assuming the result of this discussion, I shall take it as agreed that perception, as opposed to judgment, is never in error, i.e. that, whenever we perceive anything, what we perceive exists, at least so long as we are perceiving it.

If the infallibility of perception is admitted, we may apply to perception the theory of the single Objective which we found inapplicable to judgment. Take, for example, such a case as spatial relations. Suppose I see

The knife and the book: one complex perceived, then analysed.

simultaneously on my table a knife and a book, the knife being to the left of the book. Perception presents me with a complex object, consisting of the knife and the book in certain relative positions (as well as other objects, which we may ignore). If I attend to this complex object and analyze it, I can arrive at the judgment "the knife is to the left of the book". Here the knife and the book and their spatial relation are severally before my mind; but in the perception, I had the single whole "knife-to-left-of-book". Thus in perception I perceive a single complex object, while in a judgment based upon the perception I have the parts of the complex object separately though simultaneously before me. In order to perceive a complex object, such as "knife-to-left-of-book", there must be such an object, since otherwise my perception would have no object, i.e. there would not be any perceiving, since the relation of perception requires the two terms, the perceiver and the thing perceived. But if there is such an object as "knife-to-left-of-book", then the knife must be to the left of the book; hence the judgment "the knife is to the left of the book" must be true. Thus any judgment of perception, i.e. any judgment derived immediately from perception by mere analysis, must be true. (This does not enable us, in any given case, to be quite certain that such and such a judgment is true, since we may inadvertently have failed merely to analyze what was given in perception.) We see that in the case of the judgment of perception there is, corresponding to the judgment, a certain complex object which is perceived, as one complex, in the perception upon which the judgment is based. It is because there is such a complex object that the judgment is true. This complex object, in the cases where it is perceived, is the Objective of the perception. Where it is not perceived, it is still the necessary and sufficient condition of the truth of the judgment. There was such a complex event as "Charles I's death on the scaffold"; hence the judgment "Charles I died on the scaffold" is true. There never was such a complex event as "Charles I's death in his bed"; hence "Charles I died in his bed" is false. If *A* loves *B*, there is such a complex object as "*A*'s love for *B*", and vice versa; thus the existence of this complex object gives the condition for the truth of the judgment "*A* loves *B*". And the same holds in all other cases.

Judgments of perception must be true, with a caveat.

The complex as Objective of perception, and condition of truth.

A loves B against B loves A: sense, and the final definition.

We may now attempt an exact account of the "correspondence" which constitutes truth. Let us take the judgment "*A* loves *B*". This consists of a relation of the person judging to *A* and love and *B*, i.e. to the two terms *A* and *B* and the relation "love". But the judgment is not the same as the judgment "*B* loves *A*"; thus the relation must not be abstractly before the mind, but must be before it as proceeding from *A* to *B* rather than from *B* to *A*. The "corresponding" complex object which is required to make our judgment true consists of *A* related to *B* by the relation which was before us in our judgment. We may distinguish two "senses" of a relation

according as it goes from A to B or from B to A. Then the relation as it enters into the judgment must have a "sense", and in the corresponding complex it must have the same "sense". Thus the judgment that two terms have a certain relation R is a relation of the mind to the two terms and the relation R with the appropriate sense: the "corresponding" complex consists of the two terms related by the relation R with the same sense. The judgment is true when there is such a complex, and false when there is not. The same account, *mutatis mutandis*, will apply to any other judgment. This gives the definition of truth and falsehood.

Back to the datum: truth needs minds, but not the judge.

We see that, according to the above account, truth and falsehood are primarily properties of judgments, and therefore there would be no truth or falsehood if there were no minds. Nevertheless the truth or falsehood of a given judgment does not depend upon the person making it or the time when it is made, since the "corresponding" complex, upon which its truth or falsehood depends, does not contain the person judging as a constituent (except, of course, when the judgment happens to be about oneself). Thus the mixture of dependence upon mind and independence of mind, which we noticed as a characteristic of truth, is fully preserved by our theory.

What's true and how we know it come after what truth is.

The questions what things are true and what false, whether we know anything, and if so, how we come to know it, are subsequent to the question "what is truth?", and except briefly in the case of the judgment of perception, I have avoided such questions in the above discussion, not because they are of less interest, but in order to avoid confusing the issue. It is one of the reasons for the slow progress of philosophy that its fundamental questions are not, to most people, the most interesting, and therefore there is a tendency to hurry on before the foundations are secure. In order to check this tendency, it is necessary to isolate the fundamental questions, and consider them without too much regard to the later developments; and this is what, in respect of one such question, I have tried to do in the foregoing pages.