

## Reviews

John D. Caputo *Truth* (Penguin, 2013), 284 pp., £8.99

Chase Wrenn *Truth* (Polity Press, 2015), 200 pp., £14.99

Timothy M. Mosteller *Theories of Truth: An Introduction* (Bloomsbury Academic, 2014), 193 pp., \$26.99

doi:10.1017/S0031819115000261

First, the good news. John Caputo is a specialist in contemporary hermeneutics who cycles to work. His friendly and enjoyable ‘commute-length’ book (284 pages) is a guided tour of Western philosophy from Ancient Greece to Derrida, via Augustine, Descartes, Kant, Hegel, Kierkegaard and Nietzsche. English-speakers are represented only by brief but honourable mentions of J. L. Austin’s performative speech-acts and Ian Hacking’s thesis that, in order to sink our teeth into reality, ‘we need both reality and teeth’. But readers who have not read much continental philosophy, and have no plans to do so, may be interested to find that its leaders are described by Caputo (as by Sabina Lovibond in her *Ethical Formation*) as occupying various positions adjacent to those of well-known analytic philosophers.

According to Caputo, the high-water mark of Modernism (the Enlightenment) was established by Kant, who held that truth ‘is not a matter of making the mind conform to reality, but of submitting reality to the work of the mind’. The mind works by looking at what it gets from the senses and then arbitrarily inventing categories and causal sequences. Caputo comments crisply that ‘the disengagement of Kant’s pure Reason from reality, from the world in which everyone else lives, is so complete that it really does begin to look quite mad’, and that Aristotle would have ‘collapsed in laughter’ at Kant’s theories of ethics, aesthetics and love (138).

Kant’s theory was soon subjected to criticism by Hegel ‘from which it never recovered’. According to Hegel ‘Truth is not just an abstract name: it actually appears in history, sitting on the back of a horse’ as Napoleon rode into Jena at the head of his army while Hegel watched. More precisely, truth is the whole world, ‘an incessant process that never attains a fixed form’ in the way required by traditional religions and sciences with their fixed scriptures and laws. Caputo welcomes Hegel’s approach but objects that Hegel ‘remained in the grip of the deepest assumption of the enlightenment,

## Reviews

which is that the world is indeed a system of reason ... which we might think of a great German insurance company with infinite resources to underwrite all disasters and keep everything on course ... That was the sort of thing that sent Kierkegaard into spasms of ridicule' (154).

The first signs of postmodernism appear later in Kierkegaard and Nietzsche. They were bachelors who disliked democracy and the mass media and 'theme parks' (the Tivoli Gardens in Copenhagen). But they introduced laughter into philosophy, and they used it to show that truth is not discovered in an armchair or university or church but is 'grown in the wild of the real world... a matter of blood, sweat and tears' and shocks to the system.

For Kierkegaard, truth is the distinction 'between a true Christian life and a phoney one', and it turns on 'the white light of standing alone before God with all eternity hanging in the balance'. So what matters isn't to go to church on Sundays or what we say or 'the place of the history of China in the march of the [Hegelian] Absolute through time but the history of my soul, alone before God'. But while Caputo sympathises with Kierkegaard's stated desire 'to find a truth which is true *for me*, to find *the idea for which I am willing to live and to die*', he is understandably resistant to Kierkegaard's analysis of the story of Abraham and Isaac as showing that we must also be willing to *kill*: 'to counter the madness of pure reason with this kind of religious madness is no solution' (175).

While Kierkegaard attacked in the name of Christ the facile Sundays-only Christianity he saw around him, Nietzsche attacked it in the name of the anti-christ, and while Kierkegaard regarded life as the gift of divine love, Nietzsche regarded it as a piece of 'stupid cosmic luck'. Caputo says he is not sure what Nietzsche finally thought about truth, but he suggests that Nietzsche replaced religious and Enlightenment ideas of One Truth with the idea of many little truths, none of which are very important. So for Nietzsche, when we make a judgement 'the question was not whether the judgement was true or false, but what end it served – life or death ... Do you say *yes* to life – not an illusory blessed life in heaven where we don't get sick and old, but *this* life, with all its warts and bruises, all its ups and downs?' (192).

This is big-picture stuff, and so is Caputo's own theory of truth which he introduces before his accounts of Kant etc., and then develops in his final chapters. Truth is what Derrida called an 'event'. Such an event 'is not what happens but is what is going on *in* what happens ... which links us up with our guiding motif of the journey, of life as a trip whose destination is radically concealed' (76). It is also something that is not discovered but is done or made

like a promise or a confession – a result of what Augustine called *'facere veritatem'*. Caputo says 'I pin everything in my little book' on a scene in Derrida's book *Jacques Derrida* which 'repeated' (re-enacted or re-wrote) Augustine's *Confessions*: Derrida confessed to being a *religious* atheist and he prayed to his absent god 'not [in order] to go to heaven but for the future, for a radical democracy and for justice, for hospitality and forgiveness'. Caputo sees this event as the start of a postmodern faith which cannot be finalized. Change is accelerating and taking us, despite resistance from humanists, towards a 'post-human' age. Soon robots and cyborgs may become smarter than we are, and intelligent life may be set free from biology and from the Earth. The old boundaries between man and machine, material and immaterial, may disintegrate. Our minds could be preserved on hard disks and 'repeated', in terms we can't imagine, on a cosmic stage where new kinds of love and community may emerge. Perhaps even Lyotard's 'incredulity about such big stories' (104) will become a thing of the past, and perhaps it is already significant that the numeral '104' does not occur under 'Lyotard' in Caputo's index whereas some of the numerals that do occur under 'Lyotard' point to pages where Lyotard and his views are not mentioned. (Is this what used to be called a 'mistake', or is it an 'event' that playfully challenges the modern paradigm of an index and whose full implications will be revealed later?)

My only reservation about Caputo's book is that it might not be about truth, for his spectacular Zeppelin of ideas is inadvertently punctured on page 2 of Chase Wrenn's book: 'But truth is not the universe. It is a property had by the claim that chickens hatch from eggs and lacked by the claim that amphibians have fur'. In his concluding chapter, Wrenn confirms that he advocates a 'deflationary' theory of truth: 'Truth is not a property with a nature we need to explain with a philosophical theory. Rather the truth predicate is a logical device for generalization and disquotation. For a claim to be true is just for things to be as the claim says they are, and that is pretty much all there is to it' (179).

Wrenn reaches these conclusions by evaluating Epistemic theories of truth (the coherence theory and Pragmatism), Correspondence theories (Tarskian and truthmaker theories), Deflationary theories (Ramsey's redundancy theory, Quine's disquotationalism and Horwich's minimalism) and some Pluralist theories (by Crispin Wright and Michael Lynch) against four criteria:

'How plausible is the theory itself? Are there strong arguments in its favour?

## Reviews

Does the theory preserve the non-paradoxical instances of the Equivalence Schema?

Is the theory compatible with realism, and is it compatible with anti-realism?

Does the theory help to make sense of the value of truth?' (56)

Wrenn doesn't collect his results together, but they seem to be as follows. Epistemic theories fail all the tests except the last: they do say that truth is valuable. All the other theories are compatible with realism and with anti-realism too. Correspondence theories do not deliver the Equivalence Schema properly. Only Pluralist theories pass all the tests.

Now the bad news. Wrenn's opening chapter is a model of clarity and concision, but then long passages are given to the issues involved in (e.g.) spotting the tiny differences between the redundancy theory and disquotationalism (109–116). This restricts the space remaining for the multiplicity of substantial questions implied by Wrenn's survey, and some of his answers are incomplete. More disappointingly, although Wrenn's four questions look like the right ones, his discussions of them make them look quite easy and uninteresting.

Wrenn's chapter 2 explains what 'realism' means: it means that there are truths or falsehoods which can be believed but *cannot be known*, such as 'The last dinosaur died on a Wednesday' and the Continuum Hypothesis. Realism so defined is simply a consequence of the law of excluded middle, and it has no testable consequences or practical implications at all. But it's not clear why a theory of truth should also have this unattractive feature, and some of Wrenn's discussions focus instead on the idea that truth depends on what we actually know or believe (which he calls 'relativism' and then rejects).

In his chapter 3, Wrenn distinguishes (not very clearly) five kinds of value that truth might have, which he calls intrinsic value, final value, instrumental value, constitutive value and telic value. He claims that the existing arguments for assigning any of these values to truth are 'weak, except that there is good reason to think that truth is at least telically valuable' which means that 'we benefit from caring about it'. A study by Nansook Park et al. 'found that people who rate more highly on measures of curiosity, open-mindedness, critical thinking, and love of learning also report higher levels of satisfaction with their lives' (52). Wrenn points out that this evidence is no less ambiguous than any other correlation, but he accepts it and adds that caring about truth also makes it easier to achieve 'collective goods' through teamwork. 'If you and I both care about truth, we will tend to believe pretty much the same thing in the same circumstances'

– and so we will be able to locate ‘Chuck’s Café’ and meet there as agreed. Maybe so, but if the value of truth was really as lightweight as Wrenn suggests then the force of his claim that some theories of truth don’t explain its value would also be negligible. And surely something is missing here: the passions and the crusades aroused by religions and politics and science and pseudo-science and philosophy.

In the end, the issue to which Wrenn attaches most importance is his Equivalence Principle:

‘Except for those that involve paradoxes of some sort, T-biconditionals are true. An acceptable theory of truth must accommodate or explain that fact’ (9).

Wrenn argues that Epistemic and Correspondence theories are not acceptable, and that may be right (although Epistemic and Correspondence theorists don’t usually think it is right). However, Deflationary theories don’t ‘accommodate or explain’ the fact that some T-biconditionals are true either: they simply assume them (one by one, or all together). If that’s enough to pass Wrenn’s test, then it’s not a very difficult test. It seems not to go to the heart of the matter.

Does Timothy Mosteller’s book supply what is missing? It begins with an ordinary and unconditional example of truth:

‘I reach for my keys in my backpack to unlock my car. They are not there. I believe that I have left my keys on my desk. I return to my office. (The secretary lets me in for the third time this semester.) I look on my desk. I see my keys. I swipe them off the desk, mumble my apologies again to the secretary, call my wife to let her know I’ll be late (again), and head to my car.’

Mosteller comments on this ‘profound yet mundane event’ that ‘I knew truth. I experienced truth. Truth was immediately present in my experience’ (3). In that case, the next step might be to introduce a well-known truth that’s *not* immediately present in our experience, such as Galileo’s law of inertia. But instead Mosteller argues that truth is not physical and the soul and its beliefs are not material (4–7), and in the course of this argument he gives us the first of many typos (‘... while 3\* does not follow from 1\* and 3\*’: (7) – it should be ‘... from 1\* and 2\*’). He then wades through the History of Truth (sic), Coherence, Pragmatism, Deflationism, Correspondence, Phenomenology of Correspondence (Husserl), and Realism and Common Sense. The story about the keys on the desk is repeated verbatim at the start of every chapter except the

## Reviews

historical one, and is made to illustrate *all* the theories of truth under discussion.

Mosteller's book is very hard going. Although many of his points seem to me to be sound – in the sense that I agree with them when I can work out what they are – I could not understand all of his arguments or his own 'correspondence' theory (based on 'realism about universals'). One example may serve to illustrate my difficulties. Since Mosteller repeatedly asserts that 'I experienced the matching of my beliefs with my keys' (9), 'I see my belief of the keys being on the desk conforming to reality of the keys being on the desk' (10) and so on, here is Frege's 'comparison objection' to the correspondence theory, which Mosteller doesn't quote verbatim but rewrites to resemble a syllogism:

'Argument 1 (A1):

P 1.1: If two things correspond perfectly, then they are identical.

P 1.2: The correspondence theory maintains that the two things that correspond are reality and the intellect which are not identical.

C1: Therefore, two things cannot correspond perfectly.'

That's not quite what Frege intended; and to make matters worse, Mosteller then replies as follows (I omit several distracting sentences):

'In A1, the second premise (1.2) seems uncontroversial ... The premise at issue then is 1.1. Must the relation of perfect correspondence entail identity? I think that the answer is yes ... When I form a belief such as "the tile is rectangular." I form a thought which ... exemplifies the property of being *about* rectangularity and the tile on the floor exemplifies the property rectangularity. The identity between thought and object lies in the relation that both object and thought stand to the universal property of rectangularity.<sup>3</sup> Thus premise 1.1 is false. Two things can correspond perfectly because of the same universal exemplified in them without being identical particular things' (93–94. Footnote 3 adds 'For example, on Thomas Aquinas' view of truth, the soul can agree or in Frege's language "correspond perfectly" with all things ...')

It is true that, if (a)  $x$  is F, and (b) I believe that  $x$  is G, and (c)  $F = G$ , then my belief corresponds perfectly to  $x$ . But in that case, there is no 'identity between thought and object' and no 'relation [relations?]

that both object and thought stand to the universal' ... so I'm not sure what Mosteller himself was thinking when he wrote this passage.

In conclusion, Caputo's book is a delight. It magically revives questions which have been left for dead, and it avoids the varieties of madness which it describes. Whether it is about truth or not remains to be seen. Wrenn's book is short, it follows standard lines and arrives at the prevailing but disappointing Deflationary consensus (without superseding the more rigorous available texts by Ralph Kirkham and Wolfgang Künne). Mosteller's book seems not to have been proof-read, and it presents comprehension challenges that some readers may be ill-equipped to meet. None of the three books really answers Wrenn's four excellent questions. We will have to try again.

**Gary Jenkins**

[garyjenkins@phonecoop.coop](mailto:garyjenkins@phonecoop.coop)

This review first published online 22 June 2015

*Superintelligence: Paths, Dangers, Strategies*

By Nick Bostrom

Oxford University Press, Oxford, 2014, pp. xvi+328

Hardcover: \$29.95/ £18.99

ISBN: 9780199678112

doi:10.1017/S0031819115000340

Nick Bostrom's *Superintelligence* could be about a God-Machine or Frankenstein-Machine that takes control of humanity for its own perverse purposes. Or the book could be about the dawn of a new age, the historical inevitability of machines-smarter-than-humans where humans become an extinct species. Or the book could be a more realistic and less euphoric as well as a concise version of Ray Kurzweil's *The Singularity is Near: When Humans Transcend Biology* (2005). Or, *Superintelligence* could be an updated and improved version of Nietzsche's thesis of the *Superman* who trans-values all values and goes *Beyond Good and Evil*.

I think the important message of *Superintelligence* is none of the above, and is straightforward: because philosophers of morality have been unable to decide which values are ultimate, and unable to explain how values are acquired and whether values are real or not, and most crucially unable to decide on criteria for choosing which values are ultimate, we have no way of teaching very smart systems, systems smarter than humans, the goals we want them to pursue that would be congenial to humanity. Consequently, we could produce a form of superintelligence where very smart machines