



How Should We Interpret Dead Philosophers?

Tom Stern on the problem with the principle of charity

Suppose you are messaging a friend. You say: 'Let's meet up one weekend?'. She says: 'This weekend would be good! How about Saturday 4th October?'. So you look at your calendar and you notice a problem: the 4th October is not a Saturday. She either really means the 4th October (Tuesday) or she really means this Saturday (the 1st). Which one does she mean? Both contain mistakes of some kind. But my guess is that you would agree that she probably meant Saturday 1st. Why? Well, she has already agreed to meet *at the weekend*, so if you were to interpret her as suggesting Tuesday 4th, then you would be assuming she thinks that Tuesday is on the weekend. We might say that of the two options, there was a better mistake, a 'Good Thing' she might have meant (Tuesdays are at the weekend). You are more likely to get to her real meaning by interpreting her as saying the Good Thing. Put another way: she probably isn't making a foolish mistake.

Of course, in this case, what you would really have done is sent another message: '4th is a Tuesday. You mean Saturday 1st, right?'. But in the cases I'm interested in, that's not an option, because the writer is a dead philosopher and I am trying to interpret their philosophical writing. We very often have to make decisions about just what a philosopher meant. A very popular and widespread rule of interpretation proceeds exactly as you did with your friend's message. It goes as follows: if The Philosopher (equivalently: your friend) leaves it open whether they meant the Good Thing (Saturday 1st) or the Bad Thing (Tuesday 4th is at the weekend), The Interpreter (you) should say they meant the Good Thing (Saturday 1st). This is called the 'principle of charity'. Put simply: The Philosopher probably doesn't mean the Bad Thing.

Although it is very commonly used, I am suspicious of this principle. What's more, I think it is worth explaining why, because it gets to heart of what we want from the philosophers we read. But first, to give you a sense of why I don't like the principle, it will help to say something more about why I

think so many philosophical interpreters *do* like it. When I am writing about a philosopher's ideas, I often find myself confronted with two different approaches to interpreting them. First, I can think:

What interests me here is what this person really thought, even if I don't get all that much out of those ideas. It doesn't matter whether the Thing they mean is Good or Bad; what matters is that they meant it.

Call this 'reading for meaning'. Second, I can think:

What interests me is the quality of the ideas and how they relate to current philosophical debates. When choosing between the Good Thing and the Bad Thing, I'd choose the Good Thing, even if that probably isn't what the philosopher really meant.

Call this 'reading for philosophy'. The reason why I think interpreters like the principle of charity is that it makes it seem as though we don't have to choose between these two ways of reading. When you interpreted your friend's message, you were 'reading for meaning': ultimately, you wanted to know when she thought you should meet. But notice that you got to her meaning by assuming that she meant the Good Thing, which is just what we do when we 'read for philosophy'. Charity seems like a bridge between reading for philosophy and reading for meaning: we use 'good philosophy' to get to 'correct meaning', so we don't have to go for one at the expense of the other. The principle of charity assures us that reading for philosophy (i.e. reading for the Good Thing) *just is* reading for meaning.

Well, I don't think it is. My own claim is that charity makes it seem (even to The Interpreter) as though we can read for both meaning and philosophy, when in fact we are only reading for philosophy. Put simply, it's a way of seeming to be interested in what a philosopher meant, while (knowingly or not) not needing to take it into account. Often, in fact, The Interpreter is squeezing the best philosophical ideas available out of The Philosopher, perhaps including ideas that almost certainly weren't intended, while making it seem like The Interpreter is being true to The Philosopher's meaning. I'm not saying this happens every time. But it's a risk.

Why do I think this? If we were setting out some of the assumptions behind the claim that the Good Thing (reading for best philosophy) is a way to get to the meaning, we might say this:

(1) Between two rival interpretations made available by The Philosopher, we have an agreed standard by which we can decide which is the Good Thing and which is the Bad Thing.

(2) The Philosopher *meant* the Good Thing, not the Bad Thing.

Why do we need these assumptions? Without (1), we wouldn't know how to determine which Thing was Good and which Bad, so we wouldn't know how to be charitable. Without (2), charity would be misguided, at least as an attempt to give us what The Philosopher meant.

But a brief look at each assumption should give us pause for thought. Let's begin with (1). Well, do you like your philosophers to say true things? To be rational? To make claims that do not obviously contradict each other? To say original things, which you hadn't thought of before? My preference is for all of them. But sometimes you have to choose *between* them. Here's a non-philosophical example: It has been argued that in the seventeenth century, it was perfectly rational to believe in witches; it's what scientific and religious texts would have told you, what authority figures thought, and it explained all sorts of events that otherwise could not be explained. (Let's agree with this historical claim, just for the sake of my example.) It was also, of course, a false belief: there aren't witches. Now suppose someone in the seventeenth century accuses her neighbour of being a witch. There are two options for interpreting her: (a) she really believes in witches (rational, but based on a false belief); (b) she doesn't really believe in witches, so it is a false accusation (true, but not as rational for the time). Which of these options is the Good Thing and which is the Bad Thing? Would it be better to be rational and wrong, or right but less rational (and, in this case, a liar)? I'm not myself sure. More importantly, I could imagine philosophers taking a different view on

this question, because the question is itself *philosophical*: we can expect philosophers, including The Philosopher we are interpreting, to take different views on what truth, rationality, and relevance are, as well as on their relative value or importance. So it looks to me as though (1) is false, and not just in particular, tricky cases, but very often.

What about (2)? Well, (2) is in trouble already, at least in terms of how useful it is. If we can't agree on which is the Good Thing and which is the Bad, then it is difficult to see why it would help us to say that the Good Thing is the one The Philosopher means. But let's just suppose that (1) is true for a particular example: we have found an agreed standard by which some meanings are Good and some are Bad. Is (2) true? Surely not always. Doesn't everyone make mistakes? Sometimes, sadly, even the best philosopher sometimes meant the Bad Thing. While that might sound obvious, I found evidence that some interpreters do indeed assume that The Philosopher is infallible and can therefore *never* mean the Bad Thing. This makes sense if we are reading for philosophy. It makes no sense if we are reading for meaning. Recall, though, that charity was supposed to combine the two. Once we admit that (2) isn't always true, it becomes an open question, on each occasion we interpret someone, whether they meant the Good Thing or the Bad. Charity is not a reliable way to read for meaning: sometimes it gives the right answer and sometimes not.

There is more to say, both for and against the principle of charity. But I will end by saying something about why, if it operates as I have suggested, it might matter. Suppose you want to buy a book about a philosopher-let's say Mary Wollstonecraft. The bookshop has two books. The first book claims it will tell you exactly what she thought and meant. You can then decide for yourself whether what she argued was sound, original, or relevant. The second book announces that the version of her philosophy that it presents is highly relevant for some aspect of contemporary feminist theory, but that is because the book's author has adapted and modified Wollstonecraft's ideas for this century, filling in some gaps, playing down some outdated material, while keeping more or less to the spirit of her ideas. I don't know about you, but I would want to read both. Maybe you like the idea of one more than the other. That's a choice you are entitled to make. But I can't imagine *anyone* wanting to read the second book wrapped in the cover of the first: a book that claims it says what she really thought and meant, when in fact it is The Interpreter's modified, updated, and selective version. In my view, whether The Interpreter knows it or not, exactly this unwanted mix-up is what we often get via the principle of charity: The Interpreter's ideas are masquerading as The Philosopher's, perhaps without The Interpreter even realizing it. Of course, at best, what The Philosopher really thought is in fact sound and relevant, so we don't need to choose between the books on the shelf. That might even be the case with Wollstonecraft. But it won't be the case with every philosopher. How we interpret a philosophical text is, itself, something we need to be more philosophical about. Beware of convenient shortcuts.

Tom Stern is Senior Lecture in Philosophy at University College London. This essay is based on his article "Some Third Thing": Nietzsche's Words and the Principle of Charity', published in the Journal of Nietzsche Studies. His research interests include nineteenth-century German philosophy and aesthetics.

Image credit: Amelie Fontaine

Monday, 19 December 2016

Share This Story, Choose Your Platform!

Related Posts



