

If attributing a particular set of beliefs to Tom Cruise provides the best explanation of his behaviour, and allows us to predict it accurately, could it nevertheless be wrong? Discuss with reference to Dennett.

Introduction

I begin this essay by clarifying some of the terms in the question, in order to dispel any vagueness about what I am attempting to discuss. I then give an account of Dennett's idea of the "intentional stance" and explain how it might or might not work in the context of Tom Cruise. In doing this, I give some examples of how the intentional stance appears to provide explanations and predictions of Cruise's behaviour, as well as pointing out some instances where there might be some apparent problems. Having done this, I return to the question posed, before concluding that while Dennett might not be correct in all respects, the idea of the intentional stance is not one that is capable of being described as "wrong".

The “best” explanation

The question asks us to suppose that a particular approach is the “best” way to explain and predict behaviour, but I think it is worth trying to tease out exactly what this might mean.

There are two considerations that are relevant.

Firstly, when we say “best” we might mean either “best possible”, or “best available”. For example, there may be a way to preserve bananas which keeps them fresh indefinitely, and perhaps this process will one day be discovered. In practice though, the most sophisticated technique known can only guarantee freshness for, let us say, five weeks. But either of these two very different ideas might fit the description “the best way to preserve bananas”. Similarly, it may be that there is an explanation of behaviour which is “better” in every way than Dennett's, although no one has yet articulated it.

Secondly, given that alternative approaches exist, it may not be that any one of them is better than all the others *in all respects*. Perhaps freezing bananas preserves vitamins but makes them taste a bit bland, while canning them destroys most of the vitamins but retains flavour. Keeping them in a vacuum in zero gravity maintains vitamins and flavour, but costs £5,000 per banana. In terms of explaining and predicting behaviour, the simplest approach may not be the the most accurate; or perhaps one approach provides almost infallible predictions but poor explanations, while another provides very plausible explanations but unreliable predictions. When competing alternatives exist, which one is described as “best” depends on the criteria of judgement.

So for the purposes of this essay, I take the “best” explanation of Tom Cruise's behaviour to be the “best available”, rather than “best possible”; and “most convincing”, as opposed to “most accurate” or “most detailed”, or “most simply stated”, etc.. In other words, the “best explanation” in this context is the *most convincing explanation available*. I think this is in line with the spirit of Dennett's own claim: “There may someday be other strategies for attributing belief and desire and for predicting behaviour, but this is the only one we all know now” (Dennett 1989:21).

How can an explanation be wrong?

Suppose I believe that my refrigerator has a special light inside which sucks in heat and makes the inside of the refrigerator cold. Believing this, I can explain why the light appears to be on whenever I look inside, and why the casing of the light may appear warm compared to the rest of the inside of the cabinet. I could also make predictions: if I unplug the fridge, then the bulb will go out, and the fridge will start to warm up. But despite the power of this belief to explain and predict the behaviour of the fridge, we would say that it is nevertheless wrong.

There are many examples from the past of ideas which have come to be regarded as “wrong”, even though they might have been successfully used to explain and make predictions. Such ideas (such as the Ptolemaic explanation of planetary motion) may well have been regarded as the “best” explanation in one sense or another, although they have since come to be regarded as wrong.

In the same way, it could be that despite its apparent power, the *most convincing explanation available* that we have of Tom Cruise's behaviour could nevertheless be wrong: we just happen not to realise that because all the evidence we have so far supports it.

This raises the issue of *falsifiability*: are there any circumstances which we could imagine that would lead us to regard the intentional stance as being “wrong”? In the case of a scientific theory, when sufficient evidence is available to undermine it, then theory is

regarded as being “wrong” (or at least, not completely correct). Dennett may not market his idea as a “scientific theory”, and so it is perhaps not altogether fair to evaluate it on these terms, but it is nevertheless worth considering whether we could conceive of any evidence or observation that would refute the idea of the intentional stance: if no such observations can be conceived, then perhaps it is not possible that we could ever show it to be “wrong”.

In the course of this essay, I will consider what might be said against Dennett's claim, and whether this might be enough to suggest that he could be wrong.

What does Dennett say?

Dennett proposes that the way that in order to make sense of someone's behaviour, we should adopt what he calls the “intentional stance”. This means that we should treat him

as an agent, indeed a rational agent, who harbours beliefs and desires and other mental states that exhibit intentionality or “aboutness” and whose actions can be explained (or predicted) on the basis of the content of these states (Dennett 1991:76)

Dennett contrasts the intentional stance with the “physical stance”, and “design stance”, both of which he says we familiarly use to understanding the behaviour of particular systems.

In adopting the physical stance, we consider the system's constituent physical parts (down to whatever level of detail is appropriate) and then make use of our knowledge of appropriate scientific laws (mechanical, electromagnetic, quantum or whatever) to explain and predict how the system will behave under certain circumstances. We might use the physical stance, for example, to explain the behaviour of snooker balls. In principle at least, we could explain any system this way, if only we had enough information to work with.

The design stance introduces a level of abstraction, so that we predict a system's behaviour based on the assumption “that it will behave as it is designed to behave under various circumstances”. (Dennett 1989:16). For example, given that we know what a television is designed for, we can make reliable predictions as to how it will behave under certain circumstances), even though we might have little or no understanding of what physical

components it is made from, or how they work together to produce the effect that they do. Both the physical and design stance can be used to describe the behaviour of a system in a ways that are both meaningful and useful, but there are some differences worth noting:

- in adopting the design stance, we do not care how the television has been built, or how it achieves the effects it does. It may contain complex electronic circuitry or a little man with a large collection of film reels, so long as it behaves as it is designed.
- it is typically going to be a lot quicker and simpler for us to understand, describe and predict the behaviour of the television from the “design stance”

When it comes to a system which is very complex, such as Tom Cruise, the physical and design stance, though they may be useful in very limited cases, do not provide a practical means of explaining and predicting his behaviour. This is where the intentional stance comes in: by supposing that Cruise is a rational agent who harbours certain beliefs, and asking ourselves how such an agent ought to behave, we appear to have a very reliable way to explain and predict his actual behaviour. As was the case when moving from physical to design, the intentional stance allows a level of abstraction which means that by disregarding certain details, we are able to provide much more economical explanations.

For example, there is a scene in the film “Top Gun” in which Cruise sings a song. By adopting the intentional stance, it is trivial to explain Cruise's behaviour here: Cruise

desired to fulfil the demands of the script, and believed that the script called for him to sing, and so sang. We could even have predicted, some time before the event, that this is what he would do. But while we might be prepared to admit that it would theoretically be possible to explain and predict his behaviour in the same way based on a knowledge of, say, the position and motion of all atoms in and around him, such an explanation would be anything but trivial. To someone who only knows about the physical and design stances, the predications we can make about about Tom Cruise would “look like magic”, as Dennett says.

In terms of explanatory power, it would seem that the intentional stance is capable of explaining anything that Cruise does: we can appeal to Cruise's putative beliefs in accounting for his marriage to Nicole Kidman (he believed he loved her and that she loved him, among other things), his divorce from Kidman (he believed that it being divorced would ultimately make him happier) and his stated support for scientology (he believes it to be true).

All of these explanations sound very plausible.. and perhaps a little too convenient: recalling the issue of falsifiability: is there *anything* Cruise could do which could not be explained in terms of some supposed belief? While I have said that it may not be fair to treat Dennett's idea as a scientific theory, and even Dennett doesn't claim that it is infallible: “it [the intentional strategy] works with people *almost* all of the time” (Dennett 1989:21, my italics), it seems to stand up pretty well. Suppose that I hear Cruise saying “bananas

taste dreadful”. I should probably be inclined to attribute to him the belief that “bananas taste dreadful”. The next day, I see Cruise eating a banana with apparent relish, smacking his lips and saying “this tastes good”. How does the intentional stance cope with this? With very little difficulty: I simply revise the set of beliefs that I attribute to Cruise so that they accord with both observations. For example, I might still think that he believes “bananas taste dreadful”, but I now also think that he believed that it was necessary to eat a banana to provide publicity for a film. Or perhaps I say that my earlier assessment of his beliefs was wrong: he actually believes bananas taste good, but that it was in his interest to lie about it.

And even supposing we do appear to come across an instance where the intentional stance fails to provide an explanation of Cruise's behaviour, this arguably does not undermine it, because we explain the failure by supposing that Cruise is not, in Dennett's words, a “rational agent”.

In terms of its ability to predict Tom Cruise's behaviour, the intentional stance seems very impressive in its power. Katie Holmes (who I assume is fairly familiar with Tom's beliefs) may not be able to predict exactly how Tom would respond to being asked “what do you remember about our engagement?”, but it in all probability she can give a description of a small (as compared to the total number of theoretical possibilities) range of reactions which will include the correct one. And the more intimately she knows Tom's beliefs, the more narrow will be the range of reactions she will be able to predict. As already mentioned,

whatever his reaction is, she will be able to use explain it by referring to beliefs that she attributes to him (including perhaps ones that she hadn't previously been aware of).

Is the predictive power of the intentional stance in any way falsifiable? I am not sure that it is, although should Tom Cruise consistently behave in ways that are outrageously different from the predictions made by Katie Holmes then we should probably call it into question. But since any failed prediction could be retrospectively justified by appealing to some previously unknown belief, it's difficult to imagine a failed prediction which would conclusively refute the idea altogether.

So given the explanatory and predictive power of the intentional stance (and despite the caveats I have mentioned), it does not seem unreasonable to regard it as being the *most convincing explanation available* for Tom Cruise's behaviour. So could it be wrong?

Could Dennett be wrong?

As I have said, I do not think that Dennett's idea is in principle falsifiable. Does that mean they must be right? I mentioned earlier cases where we make use of ideas that have explanatory and predictive powers, but are nevertheless wrong. In fact, Dennett himself alludes to this possibility himself: “even when we are surest that the strategy works *for the wrong reasons*, it is nevertheless true that it does work, at least a little bit” (Dennett 1989:23).

In this context, Dennett is contrasting systems that behave *as if* they had beliefs and desires (those for which the strategy works “for the wrong reasons”) with systems that *really* have beliefs and desires. But the purpose of his comparison is to make the claim that there is in fact there is no difference between the two types of system: something that looks like Tom Cruise, and behaves *as if* it had beliefs and desires, *really does have* beliefs and desires. There is nothing more to a system having a belief or desire than that it so behaves.

Dennett is saying that we make a mistake if, when we attribute beliefs and desires to Tom Cruise, we regard them as being anything more than a useful fiction which helps to explain and predict his behaviour. Beliefs are no more (or less) real than the equator: they act simply as a means of expressing something in an economical and useful way.

In this context, then according to Dennett's own definition, the question of whether he could

be “wrong” takes on a slightly different flavour. Could the idea of the equator be wrong, for instance? Well, no, because the idea of the equator is a human invention, and means nothing more or less than what we define it to mean. If by saying “Tom Cruise believes that bananas taste nice” I mean exactly the same thing as if I had said “Tom Cruise behaves as he does and will behave in certain ways because he believes bananas taste nice”, then I can't be “wrong”.

If we accept Dennett's own premises then, I do not think that the intentional stance can be “wrong”. If the intentional stance is wrong, then it must be because Dennett's premises are unsound. How might this be?

Dennett identifies Tom Cruise's having a belief with his behaving as if he had one; External, visible behaviour is the only way in which “beliefs” can ever be manifested. And if this is true, then it follows that there are no such things as private beliefs. While there are ways to defend this position, it seems at odds with personal experience, where we often think of ourselves as having many beliefs and desires which are not realised in the form of external behaviour.

Perhaps we could challenge Dennett by supposing that Tom Cruise to believe that Nicole Kidman has a mole on her left shoulder-blade, a belief that will not affect his behaviour one way or another. Surely this shows that to hold a belief is not the same as to behave as if one held it? But Dennett can respond by saying that in this case Cruise *does* behave as if he had

that belief (i.e. he doesn't do anything remarkable), and our suggestion that this belief won't affect his behaviour simply confirms the predictive power of the intentional stance. So I think it is difficult to imagine being able to refute Dennett's claim with this type of thought experiment.

Perhaps we could envisage a situation where researchers discover some mechanism inside the brain which acts as a "belief repository": some kind of mechanism where intentional states are stored and can be enumerated using sophisticated medical equipment. Would this undermine Dennett's argument? Suppose we could point to something in Tom Cruise's brain and say "there's the belief that bananas taste bad". In this case, we have isolated something which is not identical to Cruise's behaving as if he has this belief (although belief and behaviour will be strongly correlated). This would mean that Dennett is wrong to describe say that having a belief as meaning the same thing as behaving as if you had it. But even if he is wrong in this respect, it seems unlikely that the intentional stance would stop working.

Conclusion

Dennett's idea of the intentional stance provides us with a very plausible explanation (albeit one which may work better with the benefit of hindsight), and what might legitimately be called the best means of predicting Tom Cruise's behaviour. Dennett does not claim that his idea is any more than a useful way of understanding the problem: it is not a scientific theory which can be falsified. Moreover, while some of Dennett's premises might be called into question, I do not think that, because of the way that the intentional stance is presented, it makes sense to describe it as “right” or “wrong”, any more than the idea of the equator could be “wrong”.

REFERENCES

- Banana (2002-2007) [WWW Document] Retrieved 26 April, 2007 from World Wide Web:
<http://en.wikipedia.org/wiki/Banana>
- Dennett, D (1981) "Brainstorms". Cambridge, Mass. MIT Press.
- Dennett, D (1989) "The Intentional Stance". Cambridge, Mass. MIT Press.
- Dennett, D (1991) "Consciousness Explained". London. Penguin Books.
- Tom Cruise (2002-2007) [WWW Document] Retrieved 26 April, 2007 from World Wide Web:
http://en.wikipedia.org/wiki/Tom_cruise