

HANDOUT

Introduction to Philosophy

True believers – Sample answers

1. What does it mean to say that someone is in a particular psychological state, for example, is in pain?

It's surprisingly hard to answer this, but there are at least four possible tacks one could take:

- Perhaps a fairly natural first thought is that to be in pain is to have a sensation that *feels* a particular way, or, to put this into philosophical jargon, to have an experience with a particular phenomenal quality. This kind of approach fits in with, though is by no means limited to, **dualism**.
- A simple **physicalist** response would say that being in pain is to be in a particular physical (probably brain) state. However, the physicalist response suffers problems as the result of the possibility of **variable realization**: it seems we can make sense of the idea of beings with quite different physiological makeup from our own suffering pain (aliens, say, or artificial intelligences). It also offers a 'hostage to fortune': when neuroscience is more developed, it may turn out that there is simply nothing in common between the brain states of each and every person who suffers pain.
- A third possibility is that to be in pain is to exhibit a set of characteristic behaviours that are associated with pain (eg saying 'ow' and avoiding the source of harm involved). This would be a **behaviourist** response.
- However, subjects can be prevented from displaying these behaviours (because of paralysis, for example, or even just a stiff upper lip). Thus we may want to say that to be in pain is not to *actually* exhibit a set of behaviours, but instead that, if someone is in pain, they *would* exhibit these behaviours as long as they were not prevented from doing so, had no reasons to do otherwise, and so on. This is no longer a simple behaviourist response because being in pain is no longer *identified* with the behaviours. However, presumably there must be *something* about the agent that makes it true to say of them that they *would* behave in certain ways if things were different. So we end up with a kind of **functionalist** position: to be in pain is to be in some state that (as long as circumstances are favourable) fulfils the role of taking a certain input (harm to the body, say) and produces certain outputs (saying 'ow!' and avoiding the source of harm, for example). It could also form the basis of a **dispositionalist** account, though as I don't think that's on the syllabus I won't discuss that further here.

2. What do we mean when we say of someone that they believe something? Does your answer to this question have anything in common with your answer to the first question? NB Don't focus too much on what Dennett says at this stage – try to think what you may have said *before* reading Dennett's account.

We could try approaches similar to those outlined above to explain what it is to be in quite a different sort of psychological state, namely that of believing that *p*.

- Perhaps to believe something is to feel a particular way when considering the **object** of the belief (namely *p*) – a sort of confidence or certainty maybe. However, this kind of approach may seem less attractive in the case of more abstract and cognitive states such as belief than it does in the case of sensations like pain, hunger and thirst.
- And perhaps it is to be in a particular brain state, but it seems even less likely that there is some one thing common to the brain states of everyone who, say, believes that Paris is the capital of France than that there is something common to the brain states of everyone who is in pain.
- We could say that to believe that *p* is to display certain behaviours, such as to assert that *p* or to act as though *p* is true. But again we run into the problem that people may be prevented from acting on their beliefs, or simply have other priorities.
- So we could instead say that to believe that *p* is to be in some state that fulfils a particular function—in this case probably something like taking in information about the world and translating it into successful action. The state may be prevented by some means from performing its proper function, but the subject can nevertheless be in such a state.

3. **Explain Dennett's account of what it is for someone to believe or desire something.**

Dennett's account probably has most in common with a functionalist or dispositionalist approach, but it is nevertheless distinct from either of those and it's important to be clear about why this is. For Dennett, to believe that p is to be *interpreted* as believing that p in something like ideal circumstances. This needs explanation. We think of fellow human beings as being (to some degree, at least!) rational. When we try to explain what they have done or are likely to do, we don't generally look for physical causes, but try to *make sense* of what they have done, or what it would make sense for them to do. This is the beginning of interpretation. When a subject, for example, walks to a shop and buys an ice cream, it makes sense of their behaviour to say they *desired* an ice cream and *believed* the shop sold ice creams. To attribute beliefs and desires to someone in this way is to *interpret* them in Dennett's sense, and this is why views like Dennett's are sometimes called **interpretationist**. Of course, people are rather more complicated than this, and we all have to balance competing desires and some of them never get acted upon, and some beliefs may never play a role in action (because they are not relevant to anything we do). This is why we need something like ideal circumstances: in order to interpret someone fully, we would need a huge amount of information about what they have done, and probably also information about what they *would* do in different circumstances. But the process Dennett proposes is the same one: with this information we should try to make sense of the subject, and attribute the beliefs and desires to them that make their actions as rational as possible.

4. **Why does Dennett think his view entails that psychological states can only be causes in an 'unspecific and unhelpful' way?**

Dennett believes that the behaviours that are the basis for interpretation are underpinned by our nature as physical beings—they are **emergent** from the physics that rules our bodies (including our brains). He seems to think that what *really* does the causing—what has **causal powers**—is the set of low level physical events that underpins the belief in question. So when we say a belief caused some action, what we must really be saying is something like: *whatever low level physical events actually underpin this belief caused the action*. This view that only low level events or properties can be *real* causes is popular but not uncontroversial. There may be other reasons for Dennett to assert that beliefs can't really be causes, but I won't go into those here.

5. **For Dennett, we can look at systems from the physical stance, the design stance or the intentional stance. Explain the difference between the three stances, and how they might lead us to different sorts of explanations. Use examples in your explanation.**

- When we take the **physical stance** towards a system, we consider it as a physicist might: what is it made of, what physical laws is it subject to, what *causal processes* is it subject to, either internally or with respect to its interaction with its environment? From the physical stance, we can predict the behaviour of a system if we know the physical laws to which it is subject (along with the starting conditions). So we might explain a stone's rolling down a hill with reference to its mass and the laws of gravity.
- When we take the **design stance** towards a system we consider what the system is for—what its purpose is. From the design stance, we can predict the behaviour of a system without knowing much, or even anything at all, about the physical laws to which it is subject. Eg, if we know a particular system is a watch, whose purpose is to tell the time, and it shows the current time as 3pm, we can predict that in an hour's time it will show the time as 4pm without having to worry at all about how the watch works, whether it is analogue or digital or anything of the sort. Note that we can also take the physical stance towards a watch, and explain its moving parts: in the case of an analogue watch how the hands turn as the result of cogs and springs and so on and hence come to be pointing in particular directions. But this explanation also misses something out—namely that what the watch is doing is telling the time. And systems that look very different from the physical stance (such as analogue and digital watches) can look similar from the design stance. Also note that whilst we can always take the physical stance towards a system, there are some systems to which it seems inappropriate to take the design stance—eg the above example of a stone rolling down a hill: in the ordinary run of events it just won't make sense to ask the *purpose* or the *function* of the stone (or the stone plus hill system).
- When we take the **intentional stance** towards a system we go yet further. We treat the system not only as having a function or purpose for others, as a watch has a purpose because it allows *us* to tell the time, but as having goals of its own, plus the capacity to form rational beliefs about the way the world is and to act rationally on those beliefs to attain its goals. We predict behaviour from the

intentional stance by asking what it would *make sense* for the system to do, given the goals it has and the input it has received (from its senses, probably).

6. **For Dennett, do you think that it is possible that someone can be interpreted as believing that p , with all the relevant information about how they do or would behave in any relevant circumstances, and yet *not* actually believe that p ? If not, why not? If so, explain how this could occur.**

I've said that Dennett is an interpretationist, but one could be interpretationist in two different ways. One could be an interpretationist about the *epistemology* of belief and desire without claiming that to believe or desire something is to be attributed as so believing or desiring in the optimal interpretation. That is, someone could argue that we *find out* what people believe or desire by interpreting them, but what it *actually* is to believe or desire is something else (such as to be in a state such that one is disposed to behave in certain ways, or to be in a particular brain state). In this case, it makes sense to suppose that even the optimal interpretation could lead us astray—we could attribute beliefs and desires on the basis of the optimal interpretation and still be wrong. Dennett's claim, however, seems to be something stronger: namely that to have particular beliefs and desires is *nothing more than* to be attributed such beliefs and desires by the optimal interpretation.

7. **People have criticized Dennett's view by saying that it has the implication that beliefs and desires are not *real* on his view. Why might they say that?**

If having particular beliefs and desires is nothing more than to be attributed such beliefs and desires by the optimal interpretation, then it can easily seem that beliefs and desires are in some way not real. This way of looking at beliefs and desires pays no attention to the intuition mentioned in questions 1 and 2 above that being in a particular mental state has something to do with *feeling* a particular way. It also seems possible that the optimal interpretation could attribute a desire or belief to a subject that that subject would under no circumstances accept that they have. According to interpretationism, the subject would in such circumstances be *wrong*, and the idea that we can be so wrong about our own mental states—that someone else could be in a better position than ourselves to say what we want and believe than we are—is somewhat counterintuitive (though not, I think, totally unbelievable). Dennett's interpretationism also makes what beliefs and desires a subject has such a matter of what other people would, or should, say about them, that it can seem that what beliefs and desires someone has doesn't really have much to do with them, or their mental state, at all. And it seems natural to think that if beliefs, for example, are real then whether or not a subject has a particular belief must be a fact about that very subject—not about what we should say about them.

8. **How does Dennett's 'Real Patterns' attempt to derail this criticism, and does his attempt work?**

We can see Dennett's position as being something like this: to have a belief that p is to display a particular *pattern* of behaviour—one that makes it appropriate to attribute the belief that p . Dennett suggests that the temptation to think that this means that beliefs therefore are not real is the result of a natural, but mistaken, view that the things that are real are either low level, or even fundamental, entities or properties or things or properties that are reducible to, or analyzable in terms of, such entities or properties. This is a complex idea, and I don't have time to go into more detail about it here. However, such a view is fairly common in philosophy. Dennett, however, gives a number of examples that suggest that we can, and in fact *do* see things—patterns—as being real even when they are not reducible or analyzable in terms of more fundamental entities or properties. If he is right, then simply that beliefs and desires are patterns of behaviour, and not reducible to facts about the physical or physiological state of the system would not be enough on its own to lead to the conclusion that beliefs and desires are not real.