

A Euthyphro Dilemma for Higher-order Theories of Consciousness*

Daniel Stoljar, ANU

1. According to a *higher-order theory* of consciousness, a psychological state of a subject is conscious if and only if the subject is conscious of being in that state. For example, suppose you are up early and catch a glimpse of a sad old fox slinking along at the end of your street. You perceive the fox. Hence you are in the state, or condition, of perceiving the fox. According to the higher-order theory, this state you are in, this state of perceiving the fox, is conscious if and only if you are conscious of perceiving the fox.

I want to ask a certain sort of question about this theory. The question is a version of what Socrates asked Euthyphro (Plato 1962, Wright 1992, Vogt Forthcoming). Euthyphro suggested, as a definition of piety, that an act is pious if and only if it is loved by the gods. Socrates asks whether an act is pious because it is loved by the gods or whether it is loved by the gods because it is pious. If Socrates were alive today he may have put it like this (Correia and Schnieder 2012, Berker 2018): is the piety of the act grounded in its being loved by the gods or is its being loved by the gods grounded in its piety?

The point is that either way Euthyphro is in trouble. If he says that the gods love the act because it is pious, he seems to have conceded, contrary to his intentions, that piety is independent of the gods' love. Perhaps the gods are good detectors of piety. Perhaps they have the good sense to love pious acts—good for them. Nevertheless, if the gods love pious acts because they are pious, the piety of pious acts is not explained by the fact that the gods love them.

Alternatively, if he says that an act is pious because it is loved by the gods, Euthyphro seems to be committing himself to an implausible consequence. Take some act we agree to be pious, e.g., donating to *Médecins sans frontières*. On the face of it this act would exist even were it not loved by the gods. After all, the gods are fickle, especially Greek gods. Maybe they love this act, maybe they don't. The problem for Euthyphro is that, if they don't love this act, it ceases to be pious. But that is implausible. If you piously give to *Médecins sans frontières*, that act remains pious whether or not some fickle god loves it.

* For comments on a previous draft I am very much indebted to Justin D'Ambrosio and Joseph Gottlieb

2. So much for Euthyphro's definition of piety, you might think. I suspect Euthyphro would take a different view. But I won't go into that here. I'm interested instead in the idea that a precisely analogous line of questioning can be put to the higher-order theory.

For a proponent of that theory offers a definition of consciousness, or at least a definition of a conscious state, in the same sense that Euthyphro offers a definition of piety. Euthyphro's theory is that what it is to be pious is to be loved by the gods. Likewise, the higher-order theory is that what it is to be a conscious state is for the subject of the state to be conscious of being in it.

If so, we may follow Socrates in asking if you are conscious of being in the state because it is conscious, or if the state is conscious because you are conscious of being in it. Admittedly, put that way the question is hard to interpret. The repeated use of the word 'conscious' makes your eyes glaze over. But there are non-equivalent versions of that word, and if we separate them out the issue comes more vividly into view.

Sometimes 'conscious' occurs in English as an adjective that applies to states. That is the notion at issue in the phrase 'conscious state.' But sometimes it is used as a transitive verb in which we speak of a person being 'conscious of' something. The usual assumption is that 'conscious of' is synonymous with, or at least very close to being synonymous with, 'aware of'. If we make that assumption, we might formulate the issue without confusing repetition, as follows. According to the higher-order theory, a psychological state of yours is conscious if and only if you are aware of being in it. And our question regarding this theory is this: are you aware of being in it because it is conscious or, alternatively, is the state conscious because you are aware of being in it?

Either way lies trouble. If we say that you are aware of being in the state because it is conscious, we seem to have agreed that the state's being conscious is independent of your awareness of being in it. Compare a case in which you become aware of a shiny object, out on the horizon. It's shiny, and you are aware of it because it is shiny. But it's not shiny because you are aware of it. Shininess in an object is a matter of how it reflects light, and this has nothing to do with your awareness. Likewise, if you are aware of your state because it is conscious, its consciousness exists independently of your being aware of it, and hence the higher-order theory provides no account of what that consciousness consists in.

Suppose alternatively your state is conscious because you are aware of being in it. That seems to have an odd consequence parallel to the one Socrates points out. On the face of it, it's possible that you are in the state and are not aware of being so. On the higher-order theory, the state would then fail to be conscious. But that's not necessarily true. When you

perceive the fox, for example, you may focus completely on its face, its scrawny sides, and the dirt on its paws. There is no doubt you are conscious of, or aware of, the fox. But are you aware in addition of perceiving the fox? Not necessarily. What you are interested in is it, not you. Hence, if your state of perceiving the fox remains conscious, as it seems to in this situation, it is not conscious because you are conscious of being in it.

3. So that is the Euthyphro dilemma for higher-order theories of consciousness. Is it any good? I think it is exceedingly good. The higher-order theory is a well-known view in philosophy of mind, and well-criticized. But even its most trenchant critics think it has versions on which it is true (Block 2011). They say, for example, that, while there is a philosophically relevant notion of consciousness for which the view is mistaken— ‘defunct’ as Block says—there are others for which it is not. I think, as against this, that if we take it as a definition, the higher-order theory is mistaken on any version; this is precisely what the Euthyphro dilemma brings out. But I also think, in a more positive vein, that working through the problem posed for the higher-order theory by the Euthyphro dilemma points in several ways toward an alternative. The aim of this paper is to substantiate these claims.

4. We might begin by looking more closely at one formulation of the higher-order theory, the *higher-order thought* theory. I have assumed so far that ‘conscious of’ is very close to ‘aware of’. But saying that doesn’t say what either expression means. On the higher-order thought theory, ‘aware’ here means something close to ‘believe’. Hence, on this view, a subject is in a conscious state if and only if they believe that they themselves are in the state. Your state of perceiving the fox is conscious if and only if you believe that you perceive the fox.

This view is a familiar one in the literature. It is defended by David Rosenthal, and by others such as Rocco Gennaro, Richard Brown, Hakwan Lau, and Joseph LeDoux; earlier versions were defended, though not under the name ‘higher-order’, by D.M. Armstrong among others (Armstrong 1968, Armstrong and Malcolm 1984, Rosenthal 2005, Lau and Rosenthal 2011, Gennaro 2012, Brown, Lau et al. 2019, LeDoux 2019). From a terminological point of view, a ‘higher-order *thought*’ theory is what you get when you take a higher-order theory and interpret the relevant notion of ‘conscious of’ in a particular way, as close to ‘believe’.

The higher-order thought theory clearly faces a Euthyphro dilemma. Do you believe you are in the state of perceiving the fox because it is conscious, or is it conscious because you believe you are in it? If you say the first, you seem to have agreed that the consciousness

that attaches to your state is independent of your belief. Again it is like believing that an object is on the horizon because it is shiny.

Alternatively, if you say the second, that the state is conscious because you believe you are in it, you need to deal with the possibility of being in the state and yet failing to believe that you are. On the higher-order thought theory, the state is in that case no longer conscious. But as before that is questionable. Suppose you are so consumed by the fox that you completely forget (and so have no beliefs about) what you are doing, at least for a short interval. On the face of it, you remain conscious of the fox, and so your state of perceiving the fox remains conscious. If so, it can't be the case that the state is conscious because you believe that you are in it. After all, you do not believe this, having temporarily forgotten completely what you are doing.

Friends of the theory may insist that you do hold the belief in question. Maybe the belief is not so demanding. Or maybe it is suppressed or inarticulate, not the sort of belief that you could formulate in words if asked. Maybe, but it doesn't matter. For even if you do believe you are in the state of perceiving the fox, it doesn't follow that this state is conscious *because* you believe this. Further, even if you do believe this, it remains as true as ever that, if you didn't, the state of perceiving would nevertheless be conscious. After all, even if you didn't believe that you are in the state of perceiving the fox, you would still focus on the fox, and so be conscious of it, as much as before.

5. There is also a different way to bring out the way in which the higher-order thought theory is vulnerable to the Euthyphro problem.

So far I've been simplifying matters. Actually, no one thinks that the *mere* fact that you believe that you are in a state makes the state conscious. Certainly Rosenthal doesn't. In principle, you could believe that you are in a state for all manner of reasons. What Rosenthal requires is that you believe you are in the state *in a particular way*. Put in terms of awareness, the idea is that the state is conscious because you are aware of being in it in a particular way.

What way exactly? Rosenthal says you must believe that you are in the state in a way that is non-perceptual and non-inferential (Rosenthal 2005). This certainly rules out some problematic cases. Suppose you believe truly that you are in a psychological state *S* because a paper you have been reading in *Linguistic Inquiry* says that you are. Perhaps *S* is the state of following some *recherché* linguistic rule. It doesn't follow that *S* is conscious. One may say without contradiction both that you believe you are in *S* and that it is unconscious. A plausible

explanation of this is that you did not arrive at the belief that you are in *S* in a way that is non-perceptual and non-inferential.

But even this is not sufficient. Suppose again you are in *S* and an amazing and unlikely thing happens. Before you even open *Linguistic Inquiry*, you get banged on the head and freakishly come to believe that you are in *S*. In this case, three things are true: you are in *S*, you believe you are in *S*, and you came to believe this in a way that is neither perceptual nor inferential. Even so it does not follow that *S* is conscious; on the contrary, it remains as unconscious as it was before.

What Rosenthal and his friends must say here, I think, is that, in order for your state to be conscious, it is required that you believe that you are in it *by introspection*. Coming to believe something by being hit on the head is not coming to believe it by introspection even if both are non-inferential and non-perceptual belief-forming processes. From this point of view, the higher-order thought theory says that a psychological state is conscious if and only if the subject of the state believes by introspection that they are in it.

But this theory also provokes a version of our question. Do you believe by introspection that you are in the state because it is conscious or is it conscious because you believe by introspection that you are in it? If you say the first, consciousness is separate from you forming the introspective belief, in the now familiar way.

What if you say the second? The problem here is that, while introspection is a non-perceptual and non-inferential process, it is also a process whereby you form beliefs about a certain class of psychological states to the effect that you are in them, namely, conscious states. If so, the higher-order thought theory is circular. It does not explain consciousness at all, but rather presupposes it. A conscious state is one that you believe that you are in in a certain way. What way? In a way that yields belief about conscious states.

You might say it's no objection to a theory that it's circular so long as it is also informative. That's true, but the issue for the higher-order thought theory is whether it is informative enough. Suppose I tell you about something you have never heard of—Xs say. You ask me what Xs are and I say, "Xs are things you detect with an X-detector." That says something about Xs; it says they are detectable, for example. But it still leaves you in the dark about what Xs are, and indeed what X-detectors are. The trouble for the higher-order thought view is that, construed as a theory of consciousness, it is uncomfortably similar to this theory of Xs.

Could a proponent of the theory try to understand introspection in a way that detaches it from consciousness? Might they say, for example, that introspection is the non-perceptual

and non-inferential process whereby you form beliefs about your psychological states—conscious or not—to the effect that you are in them? The difficulty is that it is not clear that there is any such process. At any one time you are in a huge number of psychological states, many of which are unconscious. You have various routes to the knowledge that you are in these psychological states. But, as far as the unconscious ones go, it is not plausible that you come to know that you are in them via routes that are non-inferential or non-perceptual. Unconscious states are interesting in all sorts of ways, but they are not epistemologically interesting. In fact, what is distinctive about unconscious states considered as psychological states is that you basically come to know that you are in them in exactly the way that you come to know that you are in non-psychological states.

6. We've seen that higher-order theories face the Euthyphro problem, and focused on the higher-order thought theory in particular. At this point I should emphasize that the discussion we have been having is independent of an important variation in such theories, namely, whether they assume a modest or an ambitious form.

The modest/ambitious distinction is a matter of how to interpret, not the 'conscious of' part of the higher-order theory but the 'conscious state' part (Block 2011). Any higher-order theory says that something is a conscious state if and only if you are conscious of being in the state. But what notion of a conscious state is the target of this analysis?

When people speak of conscious states they sometimes employ what philosophers call *phenomenal consciousness*, and so restrict themselves to 'phenomenally conscious states' (Nagel 1974, Block 1997, Stoljar 2016). Phenomenally conscious states are those that are essentially such that there is something it is like, or something it feels like, to be in them. The state of having pins and needles, for instance, is a paradigmatic phenomenally conscious state in this sense. There is something essentially it is like, some way it feels, to be in it.

Alternatively, when people speak of conscious states they sometimes don't restrict themselves in this way. Instead they employ a notion that applies to any psychological state at all, phenomenal or not. If you know, remember, perceive, desire, or imagine something, for example, there is a sense in which you may do these things consciously or not. Further, if you do them consciously, you are in a conscious state of knowing, remembering, perceiving etc. This way of talking does not commit us to the view that these states are phenomenally conscious states, and that remains true even if some of them are themselves phenomenally conscious states. There are various labels in the literature for consciousness in this sense: 'access consciousness,' 'introspective consciousness' and so on. Here I will call it

aphenomenal consciousness to indicate that it is a notion that applies to psychological states whether or not those states are phenomenally conscious.

From this point of view, a higher-order theory is modest if it concerns aphenomenal consciousness, and gives an account of what it is for a psychological state to be conscious in this sense. Correlatively, a higher-order theory is ambitious if it concerns phenomenal consciousness, and gives an account of what it is for a psychological state to be conscious in that sense.

7. Along with many others I think the higher-order thought theory is implausible if it is ambitious. The ambitious theory entails that the state of having pins and needles, for example, is a phenomenally conscious state—is such that there is essentially something it feels like to be in it—because you believe you are in it. That is a remarkably implausible thing to say. Admittedly, at the current stage of scientific and philosophical investigation, it is mysterious to everyone what makes a phenomenal state the state it is. But we can certainly rule out some false theories, and one of them is that you feel the way you do when you have pins and needles because you believe you are in the state of having pins and needles.

That the ambitious theory is so implausible raises some interesting interpretive questions. Does anybody really hold the theory in this form? Proponents of the higher-order thought theory often speak as if they do. But it may be that appearances are deceptive. For example, maybe they don't mean what others mean by a phenomenally conscious state. Phrases like 'what it's like' are slippery, and there are lots of opportunities for talking past one another (Gottlieb 2018).

It may also be that they aren't interested in a definition of consciousness in the sense I have been assuming. My assumption throughout has been that the higher-order theory offers a definition of a conscious state in the familiar (though admittedly elusive) sense that Euthyphro offers a definition of piety. But maybe proponents of the theory aren't after a definition like that. Maybe instead they want an account of the representational or informational structures that as a matter of fact are correlated with or realize phenomenally conscious states (Brown, Lau et al. 2019). To interpret them this way is akin to imagining Euthyphro saying to Socrates: "My dear fellow, your line of questioning is irrelevant. What made you think I'm concerned with definitions? I'm interested in what in fact happens in the world when acts are pious."

8. There is much to say about maneuvers like this in defense of the ambitious higher-order thought theory. I won't examine them here. Instead I will accept that the theory is no good as an ambitious theory, and ask about its credentials as a modest theory. For as we noted above, when people criticize the higher-order thought theory as an ambitious theory, they often assume it is fine as a modest theory. And at first sight this is reasonable enough. After all, we aren't making things up. We do have a notion of phenomenal consciousness. We do speak of states of knowledge or desire as being conscious or not, and we don't usually mean by this that they are phenomenally conscious. What then do we mean? An initially plausible answer is that we are talking about whether we believe we are in these states or not.

But the Euthyphro dilemma brings out that this initially plausible answer cannot be right. Even a modest version of the view confronts the objection. Suppose we explicitly assume that the target of the theory is states that are conscious in the phenomenal sense. Then the theory says that a psychological state is conscious in this sense if and only if you believe you are in the state in a certain way. Once again, we may ask our question. Do you believe you are in the state because it is conscious, or is the state conscious because you believe you are in it? The fact that phenomenal consciousness is in play has no bearing on the probative value of this question. If you say the first, you have conceded that consciousness even in the phenomenal sense is independent of your belief. And if you say the second, you confront again all the problems we have been looking at. In particular, if you say that the state is conscious because you believe that you are in it by introspection, you have presupposed the notion of phenomenal consciousness rather than explained it.

You might object that people are free to use words in any way they like. Some psychological states have the property of being such that their subjects believe that they are in them by introspection; some psychological states lack that property. The higher-order thought theory calls this property 'consciousness' and there is a sense in which they can't be criticized for doing so. However, while this is true, it doesn't undermine the point I am trying to make. Whatever this property is called, it remains the case that psychological states have it only if they are conscious in some prior sense. The crucial point is that consciousness in that prior sense is precisely the phenomenal notion that the modest version of the higher-order theory sets out, and fails, to capture.

9. The theme I have been emphasizing, that the Euthyphro problem undermines the higher-order thought theory in even its modest form, is surprising, as I have said. But it also raises the question of what an alternative approach to phenomenal consciousness might look

like. Here we arrive at a second theme I want to emphasize: that the points we have been making in the course of criticizing the higher-order theory show the way toward a first-order future.

According to the higher-order theory, the state of perceiving the fox is conscious if and only if you are aware of perceiving the fox in a particular way. If we wanted to formulate a first-order theory that is structurally similar, we may start with something like this: the state of perceiving the fox is conscious if and only if you are aware of what you perceive, namely the fox, in a particular way. Here the only difference is that where the higher-order theory says you must be aware of perceiving the fox, the first-order theory says that you must be aware of what you perceive.

It is easy to generalize this approach beyond perception. If you know something, for example, your state of knowledge is conscious if and only if you are aware in a certain way of what you know. Likewise, if you desire or imagine something, your state of desiring or imagining is conscious if and only if you are aware in a certain way of what you desire or imagine. Moreover, since expressions like ‘what you perceive’, ‘what you know’ etc., capture what is sometimes called the *intentional object* or the *content* of various psychological states, we might say more generally that, on the first-order theory, any psychological state is conscious if and only if the subject is aware of the intentional object or the content of the state in a particular way.

10. Since a theory like this applies to any psychological state at all, phenomenal or not, it counts as a modest rather than an ambitious theory. And since it does not entail that a conscious state is the target of some other psychological state, it is a first-order rather than a higher-order theory. There is clearly much to say about a theory like this. I won’t attempt to go into details here, though I have gone into some of them elsewhere (Stoljar 2019a, Stoljar 2019b). Instead let me make a number of points that are relevant to our discussion.

To begin with, this first-order theory says that you must be aware in a certain way of what you perceive in order that the state of perception is conscious—in what way exactly? We saw above that the higher-order thought theory answers this analogous question by appealing to introspection. In the case of the first-order theory, there are several things one might say, but the most natural is to appeal to attention, or better, since attention is a notion that comes in degree, to attention above a sufficient degree. So in the case of perceiving the fox, if that state is conscious, it will at least be true that you attend to some sufficient degree to the fox.

This appeal to attention raises a number of questions which have been the focus of quite a bit of recent literature; see, e.g., (Mole 2011, Mole, Smithies et al. 2011, Waltzl 2017, Graziano 2019, Jennings 2020). What is attention exactly and how does it relate to other psychological states and phenomena? How does the fact that attention comes in degrees affect things? How does attending to the intentional object of a state affect the functional role of the state? How does attention of this sort work in cases, such as e.g. moods (Kind 2014), that are sometimes thought not to have an intentional object in a straightforward sense?

Again, pursuing these issues would take us too far from our main topic, but there is at least the following point about attention that is relevant to the Euthyphro issue and the higher-order/first-order contrast. Sometimes psychological states make their intentional objects available as potential objects of attention, sometimes they don't. Consider again the case in which you know a *recherché* linguistic rule. It is plausible that knowing this does not make the rule a potential object of attention for you; you can't attend to the rule, as we might say, through knowing it. Nevertheless, you might attend to it in a different way, e.g., through reading about it in *Linguistic Inquiry*. This shows that the first-order theory cannot simply be, for example, that to consciously know something is to attend to what you know; what is required is that you attend to what you know through knowing it. More generally, on the first-order theory, a psychological state is conscious only if you attend to its intentional object through being in it.

If we understand it in this way, the first-order theory may offer an account of the distinction between a conscious psychological state and a non-conscious one—an important point since it is a good-making feature of the higher-order thought theory that it can so easily accommodate this distinction (Rosenthal 2008). On the first-order theory, as we have developed it to this point, a psychological state is conscious if and only if you attend to a sufficient degree to its intentional object through being in it. A slightly more explicit version of the same idea is that a psychological state is conscious if and only if (a) you are in the state and (b) you attend to a sufficient degree to its intentional object through being in it. From this point of view, the notion of an unconscious state is straightforward: a psychological state is not conscious or is unconscious if and only if (a) you are in it and (b) it is not such that you attend to a sufficient degree to its intentional object through being in it—a condition that itself may come about in different ways.

Finally—you may think I have taken too long to get to this issue—this theory does not face a version of Socrates' question to Euthyphro. Formally speaking, we can of course put a version of that question to the theory. Is the state of perceiving the fox conscious because you

attend to the fox through perceiving it, or do you attend to the fox through perceiving it because the state of perceiving it is conscious? But here the question gets no purchase; it is clear on this theory that consciousness as it attaches to the state of perceiving a fox consists in the fact that you attend to the fox through perceiving it in the way that you do.

To come at this from a different angle, in the case of the higher-order thought theory, the Euthyphro problem is generated by the fact that the element that makes the state conscious is in effect a way of responding to consciousness. That is why, when you subtract that element, you are left with a state that is conscious in precisely the sense you wanted to explain. In the case of the first-order theory, by contrast, the element that makes the state conscious is not a way of responding to consciousness; it is a way, if anything, of responding to the object or content of the underlying psychological state. Hence, when you subtract that element, you are left with a psychological state, but not one that is conscious in the relevant sense.

11. I've argued that the higher-order thought theory in even its modest form faces the Euthyphro dilemma, and have sketched a rival modest theory that does not face that problem. In a moment I'm going to turn to the final topic I want to take up: whether the same sort of dilemma confronts a different version of the higher-order theory, which I will call the higher-order acquaintance theory. But first let me deal with a source of disquiet about proceedings so far, namely, that while what I have said may be right it is also unmotivated. Aren't there many existing problems for the higher-order thought theory already? What then is the point of adding a further one?

It's true there are other objections to the higher-order thought theory, but these can be batted away. One of the things that distinguishes the Euthyphro dilemma is that it cannot be batted away.

One problem is that being conscious of something is not sufficient for it to be conscious. You might be conscious of your car keys on the kitchen bench, for example, but your car keys are not conscious. Likewise, you are conscious of the fox but that does not make the fox conscious; and that is true even if the fox is conscious.

The proponent of the higher-order thought view has an easy answer. Even when it is used as an adjective 'conscious' means something different depending on what things it applies to. A fox is conscious in one sense, the state of perceiving a fox is conscious in another. The theory is not trying to explain what it is for anything whatsoever to be conscious; rather it is explaining what it is for a psychological state to be conscious. According to the

theory, what it is for a thing of that sort to be conscious is for you to believe that you are in it in a certain way. That other things are conscious in other ways is irrelevant.

A second objection is that the higher-order thought theory misconstrues the logic of the situation. Any higher-order theory explains the consciousness of the state of perceiving the fox in terms of being aware of perceiving the fox; what the higher-order thought theory adds is that being aware of perceiving the fox is a matter of believing that you perceive the fox. But belief is a propositional state; if you believe that you perceive the fox, the belief is true if you do and false if you don't. By contrast, 'aware' as it occurs in the higher-order theory may seem non-propositional. You seem to be aware of a property, namely, the property of perceiving the fox, rather than being aware *that* such and such is the case.

Once again, the theory has an easy answer. In the context of the higher-order theory, an expression like 'you are aware of perceiving the fox' is to be interpreted on the model of 'I am aware of being in Los Angeles.' It may just be possible to construe this last sentence as reporting an abstract contemplation of the property of being in Los Angeles. But that is extremely unusual. It would normally be heard as equivalent to 'I am aware of *my* being in Los Angeles,' which in turn is equivalent to the overtly propositional, 'I am aware that I am in Los Angeles'. Hence, while 'you are aware of perceiving the fox' might have a non-propositional reading, it does not in this context.

A third problem is due to Fred Dretske (Dretske 2000). Consider a case in which you see a man with a moustache and then later see him again without one. Dretske says that, in such a case, you may be (perceptually) aware of the difference even if you don't notice the difference, and so don't believe that you are aware of the difference. The problem for the higher-order thought theory is supposed to be that your awareness of the difference is a case of perceptual consciousness, and yet you do not have the relevant higher-order belief.

Again the theory has an easy answer, namely, that while this may indeed be a case of perceiving the difference, it is not a case of consciously perceiving it. Actually, this is a straightforward prediction of the theory itself (Byrne 1997). Perceiving the difference is not conscious because the relevant higher-order belief is absent. And, in fact, something similar is true for the first-order theory I sketched above. On this view, you may perceive the difference and yet do not consciously do so because you are not aware of the difference in the right way.

What is it that leads Dretske astray? A plausible answer is his views about the relation between 'aware' and 'conscious.' Above we adopted the assumption that 'conscious of' is close to being synonymous with 'aware of.' Dretske assumes something much stronger,

namely, that ‘aware of’ is so close to ‘conscious of’ that the phrase ‘conscious awareness’ is a redundancy on the order of ‘true fact.’ This makes it impossible for him to describe the moustache case as a case of awareness that is not conscious awareness. But on both the higher-order theory and the first-order alternative I sketched, it is possible to say that there is a legitimate sense of ‘conscious of’ that is stronger than mere ‘aware of’, and this is precisely what is required to answer this objection.

The final problem I will discuss concerns the possibility of empty higher-order thoughts (Neander 1998, Block 2011, Rosenthal 2011, Weisberg 2011). On the higher-order thought theory, what makes a psychological state conscious is that you believe that you are in it in a certain way. But beliefs might be false. Suppose this one is false because it is empty: you are not in fact in the relevant psychological state at all. What is a proponent of the theory to say then? It seems bizarre to say that in this case there is a conscious state, since there is no state to be the bearer of consciousness. But it seems equally bizarre to say that in this case there no conscious state, since in this case there is something that is a sufficient condition for the state to be conscious, namely, the belief that you are in the state.

Once again, the theory has an easy answer—a point made by Wilberg but is mentioned also in Block’s discussion of this problem (Wilberg 2010, Block 2011). So far I’ve been presenting the theory as saying that a psychological state is conscious if and only if you believe you are in it. That is a common and usually adequate formulation, but it does have the disadvantage of provoking the empty higher-order thought problem. In response, a proponent of the theory may adopt a slightly different formulation, and say that a psychological state of a person is conscious if and only if (a) they are in the state and (b) they believe they are in the state. This captures the underlying idea perfectly well but also avoids the problem. On this formulation, the belief that the higher-order thought theory appeals to may be false, it can’t be false in the way that the empty higher-order thought objection requires.

It is worth noting that many proponents of the higher-order theory insist on a different response to this objection. They say the belief can be empty but that the state that is conscious exists not as such but only according to the belief, rather as certain things may exist not as such but only according to the *National Inquirer*. I won’t attempt to discuss this idea here, since it is extensively discussed elsewhere; see, e.g., (Rosenthal 2011, Weisberg 2011, Berger 2014, Brown 2015, Gottlieb 2020). But it is worth noting that interpreting the view this way has the consequence that it is no longer a definition of a conscious state in the way that it is normally taken to be, and as I have taken it to be throughout this discussion. After all, a

definition of a conscious state either is or entails something of the form ‘x is a conscious state if and only if x is...’. This entails in turn that the state that is conscious must turn up on the right-hand side of the definition. But if you say that something is a conscious state if and only if you believe such and such, and if the belief in question does not entail the existence of the relevant state, then the state does not turn up as it should on the right-hand side; hence you have not defined anything.

12. I’ve concentrated so far on one version of the higher-order theory, the higher-order thought theory, which interprets ‘aware of’ as ‘believe’. A different version, the higher-order perception theory, interprets ‘aware of’ as ‘perceive’. This theory is widely criticized in the literature and I won’t try to add to those criticisms here; see, e.g., (Shoemaker 1996). In the remainder of the paper, therefore, I want to focus on a third version of the higher-order view, which I will call the higher-order acquaintance theory.

Like any higher-order theory, this says that a subject is in a conscious psychological state if and only if the subject is conscious of being in that state. What distinguishes it is that the key notions are interpreted in a particular way. First, the notion of ‘conscious state’ is interpreted as a ‘phenomenally conscious state’ in the way introduced above. So this theory is restricted to psychological states that are such that there is essentially something it is like, or something it feels like, to be in them. Second, the notion of ‘conscious of’ is interpreted, not as ‘believe,’ but rather as ‘acquainted with,’ where in turn, this is understood (to put it rather vaguely) as a phenomenal relation of some kind that a person bears to their conscious states. Putting this together: the higher-order acquaintance theory says that a psychological state is phenomenally conscious if and only if the subject of the state is acquainted with being in it.

Several recent philosophers have defended a theory of this sort, including Brie Gertler, Michelle Montague, Galen Strawson, Martine Nida-Rümelin, David Chalmers and, in different way, Uriah Kriegel (Kriegel 2009, Chalmers 2010, Gertler 2011, Nida-Rümelin 2011, Gertler 2012, Strawson 2015, Montague 2016). There are many questions one might have about it. Here I want to concentrate on whether it faces the Euthyphro dilemma.

13. You may initially think it doesn’t. As I set it out before, the problem relied on a possibility claim. The objection to Euthyphro’s theory of piety started from the claim that it is possible that the act of donating to *Médecins sans frontières* exists and yet the gods don’t love it. Since, in such a case, the act might nevertheless be pious, piety in an act is not a matter of being loved by the gods. The parallel objection to the higher-order thought theory

likewise started from the claim that it is possible that you perceive the fox and don't believe that you do. Since, in such a case, your state of perceiving the fox might nevertheless be conscious, consciousness in a psychological state is not a matter of believing that one is in it.

I wasn't critical about this part of the reasoning before because the higher-order thought theory accepts this possibility claim. It accepts, that is, that one can perceive the fox and not believe that one does. In contrast, the higher-order acquaintance theory will not accept any counterpart claim. Take the case of pins and needles. As we saw, this is a paradigmatic phenomenally conscious state; there is something essentially it feels like to be in it. On the acquaintance theory, you are in a phenomenally conscious state if and only if you are acquainted with being in it. Is it therefore possible for you to be in the state of having pins and needles and yet not be acquainted with being in it? No; you can't be in the state of pins and needles without its being phenomenally conscious; that would be a state that does not feel like anything and there is no such state. And the acquaintance theory entails that you can't be in a phenomenally conscious state without being acquainted with being in it. In sum, the connection between phenomenal consciousness and acquaintance is on this view necessary; there is no possible case in which they come apart. Hence, one might think, the Euthyphro dilemma falls at the first hurdle.

14. However, while the higher-order acquaintance theory will reject the relevant possibility claim, it doesn't avoid the Euthyphro problem. For explanatory questions may be left open even while modal questions are not, and the Euthyphro issue is at bottom about explanation and not about possibility; this is a point well-known both in the literature on grounding and on the Euthyphro dilemma. In consequence, even if there is no possibility in which you are in the state of having pins and needles and are not acquainted with being in it, we may still ask our question. Is the state phenomenally conscious because you are acquainted with being in it, or are you acquainted with being in it because it is phenomenally conscious?

As before, either way leads to trouble. If you say the first, you have agreed that what makes pins and needles a phenomenally conscious state is that you are acquainted with being in it. But that is problematic from the point of view of the higher-order acquaintance theory. One way to bring this out is to ask whether being acquainted with having pins and needles is itself a phenomenal state or not. The usual assumption is that it is. If so, we off on a regress, which is something I have dealt with elsewhere (Stoljar 2018b). If not, the higher-order acquaintance theory entails that something non-phenomenal explains the phenomenally

conscious state of having pins and needles. But friends of the higher-order acquaintance theories will certainly resist this. They want to advance an ambitious theory in the sense described above, namely, a theory that concerns phenomenal consciousness, but they don't want a theory that is ambitious in the different sense that it attempts to explain phenomenal consciousness in terms of something else. And in any case, even if they were to advance such a theory, it is hard to see how the result would be more plausible than the ambitious version of the higher-order thought theory we looked at earlier.

What then if you say the second? In that case, phenomenal consciousness in a psychological state is to be understood as independent from acquaintance, even though they are necessarily connected. For some this is already a bridge too far: it is a necessary connection between distinct existences. But even if we refrain from playing that particular card, a theory like this at least evokes a natural question, which for at least some people will have the flavor of a rhetorical question: what is the reason for accepting that phenomenal consciousness and acquaintance necessarily go together even if they are distinct?

15. So that is the Euthyphro dilemma for the higher-order acquaintance theory. Is it any good? I think ultimately it is, but I also think the dialectical situation is more complicated for this higher-order theory than it was for the version we considered earlier. For the higher-order acquaintance theory may seem to have a ready answer to the rhetorical question just posed, namely, why acquaintance always goes together with phenomenal consciousness.

The answer starts from a well-known observation about conscious states, which is in fact lying behind the discussion we were having earlier about introspective belief being neither perceptual nor inferential. The observation is that conscious states have a distinctive epistemological character. Typically, if you are in a psychological state that is conscious, you are in a certain sort of epistemic position, namely, the position to form a justified belief to the effect that you are in the state, whether or not you do form that belief. You don't need to gather further evidence as to whether you are in the state, e.g., evidence that is based on perception or inference. Rather it is the consciousness of the state that typically puts you in this position. Contrast this with non-conscious states, whether those states are psychological or not. Suppose you know some *recherché* linguistic rule or are of a certain height. Merely being in states like this does not put you in a position to form a justified belief that you are in them; you would need to gather further evidence before you are in that sort of position.

What explains this fact? For the higher-order acquaintance theory, the answer is straightforward: it is the necessary connection between phenomenal consciousness and

acquaintance. If you are in the conscious state of having pins and needles, for example, you are on this theory acquainted with being in it, and this immediately puts you in a position to form a justified belief to the effect that you are in the state. The underlying assumption here is that, whatever acquaintance is exactly, it is at least the case that, if you are acquainted with your having some property, you immediately are in an epistemic position to form a justified belief, namely, the belief that you have that property. Hence if acquaintance is necessarily connected to consciousness, it follows that, if you are in a conscious state, you immediately have a certain sort of justification to believe that you are in it.

So at this point, the stream we have been navigating, the one that concerns the Euthyphro objection to the higher-order theory, gives way to the old man river which is the epistemology of consciousness. I am not going to try to pursue that here, which means our assessment of the Euthyphro dilemma is to an extent incomplete. What I will do, however, is to turn back and notice that our previous discussion of the higher-order *thought* theory points us toward an alternative account of the epistemological fact just noticed, an account that is not committed to acquaintance.

16. A key message of our earlier discussion is that it is a mistake to try to define consciousness in terms of introspective belief. That is precisely what the higher-order thought theory tries to do, and that is why it is vulnerable to the Euthyphro dilemma. But we should not infer that there is no necessary connection at all between consciousness and introspective belief. All that follows is that any connection here is not definitional of what it is for a psychological state to be conscious.

What then is the connection? The simplest suggestion is this: necessarily, if you are in a conscious state, you will believe that you are in it. This is quite implausible. You can consciously perceive the fox, for example, even though you don't believe that you do. For one thing, you may consciously perceive a fox and yet not be psychologically capable of forming that belief; perhaps you are an animal, for example. For another, you may consciously perceive a fox and yet not believe that you are, not because you are not capable of doing so, but because it doesn't matter to you whether you are, at least for the moment; perhaps you are extremely busy, for example.

One might try to deal with these points by offering this slightly more complicated proposal about the connection between consciousness and introspective belief: necessarily, if you are in a conscious state, you will believe that you are, provided you are psychologically capable of doing so, and provided it matters to you whether the belief is true. But even this is

insufficient. After all, it is not necessary that subjects who are in conscious states are rational, and if you are irrational enough, you may believe or fail to believe almost anything. In particular, if you are irrational enough, you may fail to believe that you are in a conscious state, even if you are, and even if the psychological conditions I just discussed are met.

This observation suggests the final account of the relation between consciousness and introspective belief I will consider, an account that has been defended by Sydney Shoemaker; I also have defended a position like this in recent work; see (Shoemaker 1996, Shoemaker 2009, Stoljar 2018a, Stoljar 2019a). On this view, the connection is articulated like this: necessarily, if you are in a conscious state, you will believe that you are, provided that you are psychologically capable of forming the relevant belief, that it matters to you that the belief is true, and that you are rational. Reformulating this slightly: if you are in a conscious state and you are a rational agent, you will be disposed to form the introspective belief that you are in the state, provided certain conditions are met.

17. Suppose now there is a necessary connection of this attenuated sort between consciousness and introspective belief. One that emerges is that we are now in a position to say something about the higher-order thought theory that is more friendly than our earlier discussion would lead you to expect, namely, there is an important grain of truth in it. Introspective belief is not definitional of consciousness, but introspective belief and consciousness go together, at least in rational agents who meet certain conditions.

Another thing is that we now have the outline of an account of the epistemic characteristic of consciousness mentioned above that is distinct from the one offered by the higher-order acquaintance theory. According to that theory, the distinctive epistemology of conscious states is explained by acquaintance. According to the rival approach, it is explained by the fact that, if one is in a conscious state, and if one is rational, one will have a disposition to form an introspective belief to the effect that one is in it, provided certain conditions are met. The belief is justified in part by the fact that it is reliably true and in part because it is the sort that rational agents in that situation form. This sort of view doesn't deny the existence of acquaintance so much as removes the reason to postulate it.

How should one decide between these proposals? One way is to focus on cases of aphenomenal consciousness. On the face of it, such cases exhibit the epistemic features of consciousness, but are not phenomenal or at any rate need not be. It seems to be as true of a conscious state of knowledge, for example, as of the state of having pins and needles, that if you are in it, you are typically in a position to form a justified belief that you are. The rival

account has no difficulty with this, since it can easily be tailored to deal with both phenomenal and aphenomenal consciousness, particularly if we rely on the first-order approach to aphenomenal consciousness I set out above. But it is far from clear that the higher-order acquaintance theory can do the same. After all, whatever acquaintance is exactly, it is associated with phenomenal rather than aphenomenal consciousness.

A second class of test cases is phenomenally conscious states that apparently lack the distinctive epistemology. I said before that typically phenomenally conscious states have that character, but it doesn't follow that they always do, still less that they necessarily do. *Prima facie*, it is possible for one to feel a certain way, and yet it be epistemologically inaccessible that one does so. Perhaps the feeling in question, while real enough, is too faint to be detectable, or too fleeting to be detectable in the moment. Or perhaps it has an inner structure unavailable to introspection; for a recent discussion of this see (Lee 2019). Or perhaps the feeling is had by a creature that can suffer but can't reason, i.e., a creature that has phenomenal states but is not rational or psychologically equipped in the way we are. It is odd to say that such a creature ought to believe it is in the state even though it is impossible for it to do so. Once again, the rival theory has no trouble with such cases, but the same is not true for the higher-order acquaintance theory. If acquaintance is necessarily connected to phenomenal consciousness, and if acquaintance necessarily has an epistemological dimension, we will have that dimension whenever we have phenomenal consciousness.

18. By themselves these remarks about the relation between the higher-order acquaintance view and the alternative decide little. While I myself endorse the alternative, I won't try to explain why in this paper. As I said, we are here at a point at which the discussion of the Euthyphro dilemma widens into a more general discussion of the epistemology of consciousness. Let me instead bring the discussion to a close by summarizing the major points I have tried to make. First, the Euthyphro dilemma raises a major problem for the higher-order theory. Second, the dilemma shows that the higher-order thought theory does not provide a definition of a conscious state of any sort, and so is no good even as a modest theory. Third, thinking through this issue provides several clues as to what a first-order modest theory might look like. Fourth, the dilemma constitutes a greater difficulty for the higher-order thought theory than do standard objections to that theory. Fifth, as regards the higher-order acquaintance theory, the dilemma brings out that the appeal to acquaintance needs to be backed up by an epistemological argument, and this is itself open to challenge. Finally, while the higher-order thought theory is mistaken, there is a grain of truth in it, and

we may exploit this in thinking about an approach to the epistemology of consciousness that is different from the one implicit in the higher-order acquaintance theory.

References

Armstrong, D. M. (1968). A Materialist Theory of the Mind. London, Routledge.

Armstrong, D. M. and N. Malcolm (1984). Consciousness and Causality: A Debate on the Nature of Mind. Oxford, Blackwell.

Berger, J. (2014). "Consciousness is not Property of States: A Reply to Willberg." Philosophical Psychology 27(6): 829-842.

Berker, S. (2018). "The Unity of Grounding." Mind 127(507): 729-777.

Block, N. (1997). On a Confusion About a Function of Consciousness. The Nature of Consciousness. N. Block, O. Flanagan and G. Güzeldere. Cambridge, MA, MIT Press.

Block, N. (2011). "The Higher Order Approach to Consciousness is Defunct." Analysis 71(3): 419-431.

Brown, R. (2015). "The HOROR Theory of Phenomenal Consciousness." Philosophical Studies 172: 1783-1794.

Brown, R., et al. (2019). "Understanding the Higher-Order Approach to Consciousness." Trends in Cognitive Sciences 23(9): 754-768.

Byrne, A. (1997). "Some Like it HOT: Consciousness and Higher-Order Thoughts." Philosophical Studies 86: 103-129.

Chalmers, D. J. (2010). The Character of Consciousness. New York, Oxford University Press.

Correia, F. and B. Schnieder, Eds. (2012). Metaphysical Grounding: Understanding the Structure of Reality. Cambridge, Cambridge University Press.

Dretske, F. (2000). Perception, Knowledge and Belief: Selected Essays. Cambridge, UK, Cambridge University Press.

Gennaro, R. J. (2012). The Consciousness Paradox: Consciousness, Concepts and Higher-order Thoughts. Cambridge, MA, MIT Press.

Gertler, B. (2011). Self-Knowledge. London, Routledge.

Gertler, B. (2012). Renewed Acquaintance. Introspection and Consciousness. D. Smithies and D. Stoljar. New York, Oxford University Press.

Gottlieb, J. (2018). "Verbal Disputes in the Theory of Consciousness." Ergo 5(12).

Gottlieb, J. (2020). "On Ambitious higher-order theories of consciousness." Philosophical Psychology 33(3): 421-441.

Graziano, M. (2019). Rethinking Consciousness: A scientific theory of subjective experience, Norton and Company.

Jennings, C. D. (2020). The Attending Mind. Cambridge, UK, Cambridge University Press.

Kind, A. (2014). The Case against Representationalism about Moods. Current Controversies in Philosophy of Mind. U. Kriegel. New York, Routledge.

Kriegel, U. (2009). Subjective Consciousness: A Self-Representational Theory. New York, Oxford University Press.

Lau, H. and D. Rosenthal (2011). "Empirical Support for Higher-order Theories of Conscious Awareness." Trends in Cognitive Sciences 15(8): 365-373.

LeDoux, J. (2019). The Deep History of Ourselves: The Four-Billion-Year Story of How We Got Conscious Brains. Canada, Penguin Random House.

Lee, A. (2019). "The Microstructure of Experience." Journal of the American Philosophical Association **5**(3): 286-305.

Mole, C. (2011). Attention Is Cognitive Unison: An Essay in Philosophical Psychology. New York, Oxford University Press.

Mole, C., et al., Eds. (2011). Attention: Philosophical and Psychological Essays. New York, Oxford University Press.

Montague, M. (2016). The Given: Experience and Its Content. Oxford, UK, Oxford University Press.

Nagel, T. (1974). "What Is It Like to Be a Bat?" The Philosophical Review **83**(4): 435-450.

Neander, K. (1998). "The Division of Phenomenal Labor: A Problem for representational theories of consciousness." Philosophical Perspectives **12**: 411-434.

Nida-Rümelin, M. (2011). "Phenomenal Presence and Perceptual Awareness: A Subjectivist Account of Perceptual Openness to the World." Philosophical Issues **21**(1): 352-383.

Plato (1962). The Collected Dialogues. Princeton, NJ, Princeton University Press.

Rosenthal, D. (2005). Consciousness and Mind. New York, Oxford University Press.

Rosenthal, D. (2008). "Consciousness and Its Function." Neuropsychologia **46**(3): 829-840.

Rosenthal, D. (2011). "Exaggerated Reports: Reply to Block." Analysis **71**(3): 431-437.

Shoemaker, S. (1996). The First-Person Perspective and Other Essays. Cambridge, Cambridge University Press.

Shoemaker, S. (2009). "Self-Intimation and Second Order Belief." Erkenntnis **71**(1): 35-51.

Stoljar, D. (2016). "The Semantics of 'What it's like' and the Nature of Consciousness." Mind **125**(500): 1161-1198.

Stoljar, D. (2018a). "Introspection and Necessity." Noûs **52**(2): 389-410.

Stoljar, D. (2018b). "The Regress Objection to Reflexive Theories of Consciousness." Analytic Philosophy **59**(3): 293-308.

Stoljar, D. (2019a). "Evans On Transparency: A Rationalist Account." Philosophical Studies **176**(8): 2067-2085.

Stoljar, D. (2019b). In Praise of Poise. Blockheads!: Essays on Ned Block's Philosophy of Mind and Consciousness. A. Pautz and D. Stoljar. Cambridge, MA, MIT Press: 511-536.

Strawson, G. (2015). "Self-intimation." Phenomenology and the Cognitive Sciences **14**(1): 1-31.

Vogt, K. M. (Forthcoming). The Euthyphro Problem Revisited. Explorations in Ethics. D. Kaspar.

Waltzl, S. (2017). Structuring Mind. Oxford, Oxford University Press.

Weisberg, J. (2011). "Misrepresenting Consciousness." Philosophical Studies **154**(3): 409-433.

Wilberg, J. (2010). "Consciousness and False HOTs." Philosophical Psychology **23**: 617-638.

Wright, C. (1992). Truth and Objectivity. Cambridge, MA, Harvard University Press.