The Externalism of Fried Duck

Thomas Bechtold

It seems that when I make a claim about what I know or want or believe, I know what it is that I know, want, or believe without my having to examine my actions or words. That is, I have some sort of authority on my own thoughts. Other people, however, must listen to what I say or observe what I do in order to have knowledge of my mental states. The idea that I know my own mental states in a way that others do not is a common intuition. Another common intuition is that our mental states are influenced by external factors, for example, when I look at a red car my mental state of wanting to drive it is influenced by the external factors that it looks to me like it would be fast. But I have no control over these external factors—and I certainly cannot know them without examining the car. It would seem that the concept of first person authority is incompatible with the notion of these external factors. In Donald Davidson's article *Knowing One's Own Mind*, he explains that this incompatibilist argument rests on two assumptions. First, that if a thought is identified by a relation to something outside the head, then the thought isn't contained in the head. Second, if a thought isn't contained in the head, it cannot be fully understood by the mind in the way required for first person authority. (102) In this paper I intend to explain the theory of self-knowledge, how incompatibilist philosophers have arrived at the above assumptions, and show how Davidson deals with the first incompatibilist assumption.

A self-knowledge theory is based on the idea that we have authority over our own thoughts—we can know them in a way that others cannot, through means that others cannot use. As Descartes said, knowledge of our own mental states is a priori,

and to have such knowledge all we need to do is reflect on our own mental activities. When John wishes to eat fried duck, he doesn't need any evidence in order to know that he wants to eat fried duck. He just thinks it because it is so—he knows his mental states without any help from evidence. Other people, on the other hand, would need some sort of evidence in order to know that John wished to eat fried duck. We would need for John to say "I want to eat fried duck." We could then say that we believe that John wants to eat fried duck, based on that evidence. Indeed, we would need the evidence in order to know what he was thinking. It appears, then, that John has some sort of connection with his own thoughts that we do not have, some sort of authority over them that allows him to know them and verify them without appealing to observable evidence.

This authority is not accessible from the second or third person perspective. Were Jane to say to John "You want to eat fried duck" (or about John "He wants to eat fried duck"), we would not assume that John wants to eat fried duck merely because Jane said that he did—we would assume that John had given some evidence to Jane that made her think that he wished to eat fried duck (if we believed what Jane said). We can see, then, that second and third person claims about a person do not have the same authority that first person claims have. It seems that sincere first person claims about thoughts[1] have an authority that claims about other person's thoughts can never have. They are somehow immediate to us. This authority is commonly called first person authority, and it is this authority that is commonly challenged by externalists.

[1]The term 'sincere' here is barring situations such as irrational thoughts, where it would seem to be difficult to say that John would know everything about his thoughts or to say that he is in a position of knowing their veracity.

There are a number of different ways to illustrate externalism through thought experiments. Tyler Burge uses the example of a person on earth (I'll call her Jane) who thinks that she has arthritis in her thigh. On earth, the term 'arthritis' means a rheumatoid illness affecting only the joints, not the thighs, and so Jane would be wrong in her arthritis thoughts if she were on earth. In another world lives Jane's physical double, Twin Jane, who also believes that she has arthritis in her thigh. On this other world, however, it is the case that the term 'arthritis' refers to rheumatoid illnesses in the joints and in the thigh. Both women are in the exact same physical state, but one of them is right about her thoughts and the other one is wrong. It would seem that in order for either of the twins to know if what they were thinking is correct or not they would have to know certain things about their community by investigating certain facts about the outside world. If we truly had authoritative self-knowledge we should be able to distinguish arthritis from twin arthritis without difficulty through reflecting on our mental activities as Descartes suggested. Instead, it seems that we could not know whether we were having arthritis thoughts or twin arthritis thoughts unless we knew which earth we were on—by investigating something external.

Burge's thought experiment succeeds in illustrating one possible complaint of externalists when faced with first person authority. If we must examine the external world in order to know our mental states, we are not in a position of authority in regards to those mental states. According to externalists, then, John does not have authority over his thoughts of wishing to eat fried duck. In order to know that he wants to eat fried duck John would have to investigate the world he was in to find out what fried duck was in the first case. Let us say that John learned the taste of fried duck from eating it once in a restaurant last year. His concept of the taste of fried duck comes in some way from the fried duck that he ate at a restaurant last year.

He has already investigated the external world to know something about his thoughts (when he first experienced the taste of fried duck), but let us take it a step further to demonstrate how John must continue to rely on external factors in his claim that he wishes to eat fried duck. There are many ways of frying a duck, and each way can have a different taste. Since John does not have a complete idea of what fried duck is from his limited exposure (he doesn't know all the different tastes of fried duck), he cannot be in a state of complete authority over his thoughts of wanting to eat it. If John believes that he wants to eat fried duck, it is likely that he wishes to have the same taste experience that he had when he ate fried duck at the restaurant. It is possible that John would despise other preparations of fried duck, and when his thought is applied to those preparations of the dish his thoughts would be incorrect—he would not want to eat that fried duck. His knowledge of fried duck is causally related to something that is outside his mind, something that he cannot know everything about, and thus his belief that he wishes to eat it is not something that he can have authority over.

This situation demonstrates another claim that externalists often make about knowledge, that of learning. According to some externalists, if an idea is reached through some causal connection to the outside world, if a concept is learned through investigation of evidence rather than through introspection or reflection on ones own thoughts, then any thought that relates to that concept is also not an example of authoritative first person knowledge. This is the first assumption from Davidson's article that I mentioned above; if a thought is identified by a relation to something outside the head, it isn't wholly in the head.

Davidson says that it “doesn't follow, simply from the fact that meanings are identified in part by relations to objects outside the head, that meanings aren't in

the head." He uses the example of sunburned skin to illustrate this statement, saying that just because being sunburned presupposes the existence of the sun, it doesn't mean that the sunburn is not a condition of the skin. Let us say that Jane has a sunburn and Twin Jane has a skin condition that is physically identical to sunburn, but is not sunburn. Their respective skin conditions are identical in the physical sense, but one of them came about through too much exposure to the sun and the other did not. In other words, there are no physical differences between their skins. The only difference between the two skin conditions is that one came about through a causal connection to the sun and the other did not. Yet both are still conditions of the skin.

Similarly, consider two identical mental states. John's mental state came about through a causal connection to some external factor, and Twin John's came about through self-reflection. According to Davidson's sunburn analogy, we cannot say that one is a state of the mind and the other is not. Davidson says that this is sufficient to demonstrate that having an "appreciation of the external factors that enter into our common ways of identifying mental states does not discredit an identity theory of the mental and the physical." In other words, even though external factors enter into our mental states, these states are still in and of the mind. In the arthritis example, Jane would examine her mental states *in light of her surroundings* and she would discover that she was wrong. This doesn't mean that she was not in an authoritative position in regards to her mental states, it just means that she was mistaken about their correlation to the outside world.

The first assumption of incompatibilism is not accurate, then. A thought can be identified by a relation to something outside the head and still be entirely in the head—John's thoughts about fried duck are still in his head, even though he must find out the taste of fried duck from external factors. Of course, it remains for Davidson to

show the second assumption to be incorrect or faulty, but that is the topic of a different paper.