Robert Durfee

Alex Byrne, David Balcarras

24.00 – Topic III

April 10, 2018

**Circular Theory-Theory**

Sarah Paul argues for her theory-theory of self-knowledge in the paper “John Doe and Richard Roe.” This theory rejects the premises of the Cartesian and inner-sense theories of self-knowledge in favor of a modified Rylean argument. With this argument, Paul seems to come to a satisfactory, economical conclusion as to why we sometimes fail to correctly know our own mental state yet seem to have more insight to our own attitudes and emotions than others’. In this paper, I will outline Paul’s theory-theory and how she’s justified in rejecting the Cartesian and inner-sense premises. However, I will continue by showing how the theory-theory results in a circular method for ascribing mental states to oneself where one’s interpretation of evidence is filtered through one’s current mental state.

To be more explicit, Paul begins by listing the following claims which constitute the Cartesian, or “common-wisdom,” theory for self-knowledge:

[No Evidence]: When a subject ascribes a mental state *M* to herself, it is not on the basis

of evidence or grounds for thinking that she is in *M*.

[Infallible]: If a subject sincerely believes herself to be in some mental state *M*, then she is

in *M*.

[First-Personal]: The method a subject normally uses to ascribe mental states to herself

cannot be used in the same way to ascribe mental states to other people.

Although at first glance all these claims seem obviously true, Paul introduces a scenario, taken from Ludwig Wittgenstein, which brings this into question. Imagine you, and everyone around you, has been given a box. Everyone is instructed to refer to the contents of the box as a “beetle,” but the contents may or may not actually be a beetle. Since you cannot actually look into another’s box, you have no way of knowing what “beetle” means to anyone else.

Thets

Asdf