Insofar as we have good reason to suspect that human minds besides our own are conscious, and insofar as we have good reason to suspect that even very advanced and complex supercomputers – whatever else they can do – are not conscious, Searle is on good sturdy ground. And, insofar as Searle contends that it is more intuitively correct to associate mind with consciousness (which seems almost a cognate) and intentionality, rather than something as abstract and overwrought as rule-governed symbol manipulation systems, he seems to have the high ground. However, one can grant all this, and still defend a minimal functionalism like the one outlined in section 2 above. One simply has to concede that the minimum functionalism one is defending is not a full-blown theory of mind, but rather a research programme in some properly cognitive science. What this amounts to is a denial of the conflation of a cognitive system and a mind. It would require the view that minds instantiate cognitive systems – defined as rule-governed symbol manipulation systems – along with a lot of other things (like emotions, consciousness, intentions, judgments, etc.); and also that cognitive systems are instantiated by minds along with a lot of other things (digital computers, Martians, Super-Monkeys, etc.) This does not seem like an unintuitive step to take.

The relevant competition then becomes not some other theory of what minds are, but the completely different claim that the correct way to study cognitive systems is to study the hardware that instantiates cognitive systems. If we are going to begin from the perspective that human brains are the archetypal or representative hardware that best exemplifies cognitive systems, then such a competing claim would be some form of biological reductionism. If we are going to begin from the perspective that digital

computers are the best exemplars of cognitive systems, then some sort of computer science (where the physical-electronic structure of processor chips was the focus) would seem appropriate. The functionalist claim would be that though both of these studies would be valuable to a complete cognitive science, the crucial focus of such a complete cognitive science would be on the things that both of those physical structures instantiate – cognitive systems and processes, where cognitive systems is defined as a system in which symbols are manipulated according to rules, regardless of whether the system in question has a mind or not.

* * *

It may at first seem that a cognitive science that is not a science of the human mind *per se* is not really a cognitive science at all, or at least a valuable cognitive science. In order to defend this view, it may be helpful to revisit the Turing Test and to look at one of its interesting consequences. Recall exactly how it is that a Turing machine can pass the Turing Test; it must be able to fool a human interlocutor into thinking that it is itself a human. A digital computer (which is a powerful Turing machine) is capable of being programmed such that it can execute perfect mathematical operations and grammatical language usage time after time after time. But if we programmed such a digital computer to perform at peak competence, acting as a perfect cognitive system, it would paradoxically be a dismal failure in the Turing Test. It would be easy to determine that it was in fact a Turing machine and not a human because a human would make statistically a significantly higher number of errors than a machine would. So, in order for a machine to pass the Turing test, it would have to be programmed to make the same mistakes with the same frequency that a human would. That is, in order for it to pass our demanding

standards to prove itself a truly cognitive system, it would have to be as dumb as we are. But surely this is odd. If a functionalist cognitive system is one in which symbols are manipulated according to rules, it doesn't seem right that an exemplary cognitive system (the one against which other candidates for cognitive systems are judged) is one that routinely manipulates symbols according to rules badly. What could be the rub?

The rub, it seems, is exactly what is at issue in distinguishing a cognitive system from a mind, especially a human mind. The reason that we get this paradoxical result is because we implicitly conflate being a good cognitive system with being (or being part of) a mind. Since the human mind is our first and most well known example of something that instantiates a cognitive system, it tends to be the standard by which we judge or identify other cognitive systems. This being the case, the tendency to resort to biological reductionism in the study of cognitive systems becomes more understandable. If we come to the table with the prejudice that the human mind is the paradigm cognitive system, then we will (with the paradoxical conclusion just illustrated) judge other cognitive systems by the degree to which they resemble the human mind. Accordingly, we will judge other cognitive platforms or hardware against the hardware that instantiates human minds, namely human brains. So, even if we are sympathetic to the functionalist in allowing that cognitive systems may be multiply instantiable, we may still fall into biological reductionism by assuming that the cognitive system instantiated by the human brain has a privileged place in our cognitive science.

The result seems to be two very different research programmes with two very different assumptions:

(RP1) The minimal functionalist programme – We want to understand cognitive systems, where cognitive systems are multiply instantiable systems that manipulate symbols according to rules. Human brains are one kind of platform that instantiates such cognitive systems, and as such, their study may aid in the study of cognitive systems.

and

(RP2) The functionalist biological reductionist programme[1] – We want to understand human minds and brains. Cognitive systems are one of the things that human brains instantiate. The study of cognitive systems instantiated by other platforms may aid in the study of the human mind/brain.

One need not deny either of these agendas in order to defend the importance of the other. But because of the anthropocentric bias revealed by the Turing Test, (RP1) may require special argumentation for its importance and relevance as a research programme in cognitive science. Fodor, among others, has attempted to defend such a research programme throughout his career. In the next section, I will analyze and defend one of his principal arguments in favor of this kind of cognitive science, and I will try to provide further arguments in its support.

[1] A non-functionalist biological reductionist program, I take it, would deny the part about multiple instantiability.