

The Chinese room argument

PHIL 20208

Jeff Speaks

September 26, 2006

Most of the arguments we have considered against functionalism so far have focused on mental states associated with ‘qualia’: perceptual experiences, like visual experiences of red, and bodily sensations, like pains. But in “Minds, Brains, and Persons,” John Searle gave an argument against functionalism applicable to the (apparently non-qualia-involving) mental state of *understanding*:

“One way to test any theory of the mind is to ask oneself what it would be like if my mind actually worked on the principles that the theory says all minds work on. . . . Suppose that I am locked in a room and given a large batch of Chinese writing. Suppose furthermore (as is indeed the case) that I know no Chinese, either written or spoken, and that I’m not even confident that I could distinguish Chinese writing from, say Japanese writing or meaningless squiggles. . . . Now suppose further that after this first batch of Chinese writing I am given a second batch of Chinese script together with a set of rules for correlating the second batch with the first batch. The rules are in English, and I understand the rules as well as any other native speaker of English. They enable me to correlate one set of . . . symbols . . . with another. . . . Now suppose also that I am given a third batch of Chinese symbols together with some instructions, again in English, that enable me to correlate elements of this third batch with the first two batches, and these rules instruct me how to give back certain Chinese symbols with certain sorts of shapes in response to certain sorts of shapes given me in the third batch. Unknown to me, the people who are giving me all of these symbols call the first batch a ‘script’, they call the second batch a ‘story’, and they call the third batch ‘questions.’ Furthermore, they call the symbols I give them back in response to the third batch ‘answers to the questions’, and the set of rules in English they gave me, they call ‘the program.’ . . . Suppose also that after a while I get so good at following the instructions for manipulating the Chinese symbols and the programmers get so good at writing the program that from the external point of view — that is, from the point of view of somebody outside the room in which I am locked — my answers to the questions are absolutely indistinguishable from those of native Chinese speakers. . . .

. . . As regards the first claim, it seems to me quite obvious in the example that I do not understand a word of the Chinese stories. I have inputs and outputs that are indistinguishable from those of the native Chinese speaker, . . . but I still understand nothing.”

Do you agree with Searle that the man in the room understands no Chinese? What, if anything, does this show about functionalism? How should the functionalist respond?