Response

Thomas Eudaly
University of Illinois

Searle's Chinese room argument deals with whether the instantiation of a program for understanding is sufficient for the realization of intentional states. Intentional states include such states as believing, thinking, understanding, and so on. Typically, such states are expressible by statements with a "that p" clause, where p is a complete sentence. For example, "I believe that it is raining" expresses the intentional state which obtains when the person uttering it believes that it is raining.

Now, if one is a materialist, one thinks that mental states are brain events. Intentional states on this view, then, are realized by brains. How are brains able to do this? Because they have the required causal properties. However, if one is a computer functionalist, one thinks that intentional states are functional states realized by following a program. By "following a program," I mean taking data as inputs, following instructions for the manipulation of those data, and producing output as the result of that manipulation.

On this view, the mind is to the brain as the software of a computer is to its hardware. For example, understanding which is an intentional state, obtains if a _program_ of sufficient complexity is realized. It is the program, not the hardware, then, which is responsible for the understanding achieved by brains. Thus, we could produce a black box of unspecified construction which, if it realized the same program that a brain realizes when it understands, would have the same intentional states as a brain. "S understands p" (where, say, p is a story) means, then, that, given p as input, S realizes a program which enables S to produce output which is absolutely indistinguishable from that of a native speaker of the language is which p is encoded.[1]

The point of Searle's argument is to produce a counter-example in which something realizes a program without understanding. The man in the room is, say, given a Chinese translation of the poem, "Twinkle, twinkle little star" as input. By following the instruction manual, he is able to answer the queries of his Chinese interlocutors in a way indistinguishable

211

from that of a native speaker of Chinese. Yet, he does not understand the poem, the questions or his answers to the questions. Thus, understanding is not simply a matter of realizing a program. This is Searle's negative thesis. His positive thesis is that understanding results from the operation of certain bio-chemical devices, brains, with the causal powers necessary for producing understanding. Intentionality, then, is due not only to software but also to hardware.

This, then, is Searle's argument. I find it partially convincing. That is, I think Searle has shown conclusively that the Turing test is an inadequate test of understanding. However, I do not think Searle argues adequately for his positive thesis. Therefore, I will only challenge Searle's positive thesis and Whitmer's "causal argument," which purports to fill in its details.

Whitmer's causal argument is that, since brains have intentional states in virtue of their causal powers, a computer brain which has those causal powers will have intentional states. For a computer brain to have these powers, its components must function in the same or in an analogous way to the components of a human brain. Now, the brain functions electrochemically. Therefore, if something has intentional states, then it accomplishes this electro-chemically.

The basic problem with Whitmer's argument is that it is invalid. That is, the conclusion does not follow from the premises even if those premises are true. From the fact that the causal powers of the brain are sufficient to produce intentional states, we cannot infer that the causal powers of the brain are necessary to produce intentional states, that only those powers, whatever they are, are able to do so. Thus, all that Whitmer can conclude is that if something artificial functions in the electro-chemical way that brains do, then it has intentional states.

Now, Whitmer does state the conclusion aright where he says that "it seems conceivable that an entity could exist whose component mental parts interact magnetically or optically, and we would still want to claim that if they had the appropriate causal powers, they would also have intentionality."[2] However, Whitmer seems to trade on a confusion between the valid and invalid inference. Thus, Whitmer says that "the causal argument based on Searle's various claims admits what it seems reasonable to admit to the class of 'things with intentional states' while excluding all those things that were shown to be inadmissable by virtue of Searle's demon."[3]

That is, Whitmer wants to argue that if something is not constructed along the lines of a human brain, then it lacks the causal powers of the brain. Lacking those causal powers, it may be excluded from the class of things with intentional states. However, for this

last proposition to hold, the invalid conclusion, which is its contrapositive, would have had to have been shown.

Basically, the confusion lies in thinking that Searle's argument entails his positive thesis. Rather, they are entirely separate. Thus, although Searle says in "Minds, Brains and Programs" that "the computer understanding is not just (like my understanding of German) partial or incomplete: it is zero."[4] This doesn't <u>follow</u> from his argument. What follows is that we cannot infer that they do understand anything from the fact that they run programs. Searle, however, seems to think that his argument implies more than it does. He wants to infer from it that only brain-like things can understand. But, this doesn't follow and he offers nothing else in support of it. Whitmer's causal argument, then, simply places Searle's mistake into bold relief. From the fact that digital computers don't think in virtue of running programs, we cannot even infer that digital computers don't think, let alone infer that only things with the causal powers of the brain can think.

.

NOTES

[1]This is Schank's version of the Turing test. See Schank and Abelson, <u>Scripts</u> <u>Plan</u> <u>Goals</u> <u>and</u> <u>Understanding</u> (John Wiley and Sons, 1977).

[2]J. Whitmer, "Intentionality, Artificial Intelligence and the Causal Powers of the Brain," p. 204 above.

[3]<u>Ibid</u>.

[4]<u>The</u> <u>Behavioral</u> <u>and</u> <u>Brain</u> <u>Sciences</u>, 3 (1980), p. 419.