

## CHINESE ROOM PROJECT OVERVIEW

John Rogers Searle (July 31, 1932 – September 17, 2025) was an American philosopher widely noted for contributions to the philosophy of language, philosophy of mind, and social philosophy.

He is best known for his Chinese Room thought experiment, which challenges the notion of a truly intelligent artificial intelligence (AI). In it, he imagines a room containing an individual, who speaks only English, working with a set of instructions in English language (a “guidebook”) that allows the individual inside the Chinese Room to assemble a series of Chinese characters for anonymous communication with Chinese speakers outside the room. If the individual in the Chinese Room follows the guidebook carefully, Chinese speakers outside the room can be fooled into thinking they’re talking to someone (or something) who understands their language. We can posit that the system as a whole does in fact have something akin to an understanding of Chinese, though the individual inside the room does not.

The Chinese Room argument holds that a computer executing a program cannot have a mind, understanding, or a consciousness, regardless of how intelligently or human-like the program may make the computer behave. Searle’s Chinese Room suggests deep complications as to whether one can truly describe convincing simulations of intelligence as being intelligent. The scenario involves a human English speaker enclosed in a room with Chinese symbols and a rule book for manipulating them. Unbeknownst to the person in the Chinese Room, the symbols are Chinese characters, and the input he or she receives arrive in the form of questions. Nevertheless, by following the rules, the person inside the Chinese Room produces correct answers and so, appears intelligent to outside observers.

In one sense, Searle’s thought experiment parodies the Turing test, a test for artificial intelligence proposed by Alan Turing in 1950, and echoing René Descartes’ ideas about distinguishing thinking souls from unthinking automata.

We can note that Turing himself offered no metaphysical guesses as to what thought is in the first place. On the one hand, an observer might conclude that computers really do, or eventually will “think” and, on the other hand, there are suggestions that thought itself is essentially a form of computation. One point of view is that even if thought is not essentially just computation, computers (even present-day ones), nevertheless, might really think. Searle’s Chinese Room contrasts this notion with the proposition that computers only simulate thought and that their seeming understanding is not real understanding.

Whether or not today’s Large Language Models (LLMs) are simply stochastic parrots, the fact that AI behaviors seem to imply thought has propelled assorted hysterical notions about AI, including numerous Hollywood productions about doomsday AIs taking over the world. If Searle’s argument merely challenges theoretic or metaphysical identification of thought with computation, then the behavioral evidence – and consequently Turing’s point – remains undisputed. Since computers seem, on the face of things, to think, the conclusion that essential nonidentity of thought with computation would seem to warrant that whatever else thought essentially is, computers have it too; not, as Searle maintains, that computers’ seeming thought-like performances are bogus.

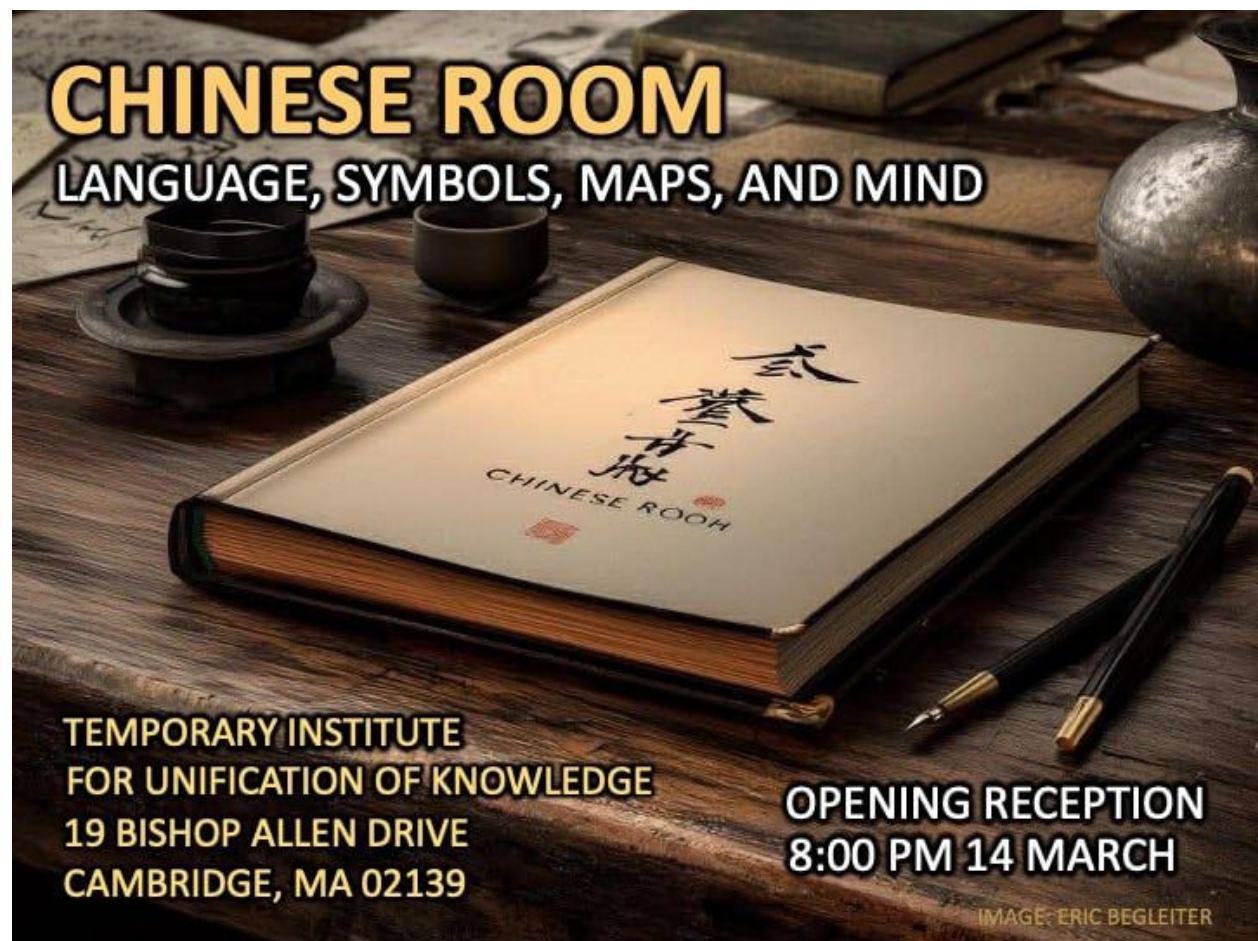
Searle never actually constructed his Chinese Room experiment in the real world. Fellows of Temporary Institute for Unification of Knowledge have now done so – partly as an experiment to investigate how a fundamental “guidebook” could be created, and further, to investigate how the

behaviors of various participants in our project may or may not correspond with Searle's initial premise.

We realize that our real-world experiment may be innately flawed, and that, at the very least, it may help to suggest how future such experiments might be conducted. We recognize too that this experiment will not resolve the many important questions about potential cognitive aspects of AI, and like controversy about what is or, is not "thought," theorists have yet to conclusively decide what "intelligence" is in the first place.

---

TEMPORARY INSTITUTE FOR UNIFICATION OF KNOWLEDGE WISHES TO RECOGNIZE THE CONTRIBUTIONS OF OUR COLLABORATORS INCLUDING MARTY ZHOU, ECHO LIU, JOE DAVIS, TREY PECK, SONIA COOK-BROEN, JOE BENWAY, DIANA KENNEY, AND DIMITRI IOANNIDIS.



**CHINESE ROOM**  
**LANGUAGE, SYMBOLS, MAPS, AND MIND**

TEMPORARY INSTITUTE  
FOR UNIFICATION OF KNOWLEDGE  
19 BISHOP ALLEN DRIVE  
CAMBRIDGE, MA 02139

**OPENING RECEPTION**  
**8:00 PM 14 MARCH**

IMAGE: ERIC BEGLEITER