

The Turing Betrayal

From Comparative Mimicry Heuristic to the Ideology of Human Replacement

Emmanuele Somma ¹

¹Epistechonology Research Project

Abstract

Turing's [imitation game](<https://plato.stanford.edu/entries/turing-test/>) became a shorthand for machine intelligence. Its reception often collapses a question about observable performance into claims about understanding, then into a design preference for replacing people.

Keywords: *turing, artificial intelligence, imitation game, epistemology, intelligence augmentation*

Alan Turing asked whether machines can think in 1950, then set the question aside. He proposed a conversational contest using written slips between an interrogator and unseen participants. The shift was operational: replace metaphysical definitions of mind with comparative performance in a controlled setup. It never defined what thought is.

Popular discourse turned that heuristic into an engineering dogma: if a machine passes for human in conversation, it is intelligent. A second step treats fluent mimicry as proof of semantic understanding. A third promotes human replacement into an institutional imperative.

This progression is *the Turing betrayal*. It marks the historical reception of Turing's text, not an intentional plot by Turing himself.

1 The Bifurcation in the Game

The opening test in "[Computing Machinery and Intelligence](#)" has three players: a man, a woman, and an interrogator. The interrogator attempts to distinguish between them through typed answers. Turing asks what happens when a machine takes the man's place, and whether the interrogator errs as frequently as in the gender game (pp. 433-434). The comparison evaluates performance within an existing social deception.

Later in the essay, Turing quietly introduces a different scenario: a machine and a human being questioned directly, with the interrogator attempting to tell them apart. This second version is what the world remembers as the Turing Test.

Susan Sterrett demonstrated the divergence between these formulations in "[Turing's Two Tests for Intelligence](#)". In the first version, a machine could outperform a human player at a specific communicative task. In the second, suc-

cess is purely negative: the interrogator's failure to distinguish synthetic replies from human text (pp. 541-545). Anthropomorphism becomes the sole benchmark.

Sterrett raises the essential historical caveat: Turing treated both formulations as practically equivalent. The ambiguity originates in his own pages. Deception was not added by careless commentators; deception was built into the game's architecture.

2 Performance vs. Explanation

An interrogator's error is a raw empirical observation. It registers a local event: a specific judge, in a specific exchange, under specific constraints. Inferring that the machine possesses comprehension requires a separate, ungrounded leap.

John Searle challenged this leap in "[Minds, Brains, and Programs](#)". The [Chinese room argument](#) shows that manipulating symbols according to formal rules does not constitute linguistic understanding (1980, pp. 417-424). Regardless of objections to Searle's thought experiment, its core point stands: behavioral performance cannot serve as its own explanation.

Hector Levesque analyzed the practical consequences of this substitution in "[On Our Best Behaviour](#)". When a benchmark rewards the outward appearance of intelligence, engineering optimizes for cheap theatricality and conversational evasions. The critical inquiry asks what underlying mechanisms generate the answers, how they function, and where their failure modes lie. Fluent syntax is a symptom to investigate, never an explanation.

Defenders of the behavioral test reply that we judge [other human minds](#) solely by observable conduct. We cannot inspect another person's consciousness directly. Demanding privileged access to inner subjective states would

EIN: 5 Revision: 1 Broadcast: Episode 5

Canonical URL: <https://www.epistechonology.org/the-turing-betrayal/>

Audio Edition: <https://www.epistechonology.org/podcast/epistech.005.005.001.the-turing-betrayal.mp3>

Corresponding author: emmanuele@epistechonology.org

This work is licensed under [Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International \(CC BY-NC-ND 4.0\)](#).

disqualify humans as well. The reply misses the target. Epistechology does not demand metaphysical access to inner minds. It rejects the inflation of a constrained conversational score into a complete theory of intelligence.

3 The Learning Machine

Turing's paper resists fixation on finished impersonators. Its concluding section explores learning machines, proposing a "child-machine" educated through interaction rather than programmed with static encyclopedias (pp. 455-459). Bernardo Gonçalves argues in "[The Turing Test is a Thought Experiment](#)" that the imitation game was designed to destabilize anthropocentric biases and open research into machine learning.

The paper is an unstable experimental brief: a provocative parlor game, divergent setups, and serious reflections on cognitive development. The betrayal occurs when institutional forces strip away the instability and convert the easiest deceptive score into an immutable standard.

4 The Economics of Imitation

The consequences of this reduction are economic and political. Erik Brynjolfsson analyzes the distortions of "[The Turing Trap](#)": corporate investment overwhelmingly funds systems designed to automate and replace human labor, neglecting technologies that augment human capability (2022, pp. 272-287).

When the cultural measure of technological progress is how convincingly an algorithm impersonates a person, labor displacement appears as natural engineering destiny. A different benchmark would evaluate whether a tool assists an investigator in diagnosing illness, helps a citizen audit a bureaucratic decision, or enables an artisan to master a difficult craft. These are architectural choices. Neither follows inevitably from the concept of intelligence.

Kurt Vonnegut anticipated this systemic inversion in *Player Piano* (1952). His automated society does not merely replace physical workers with automated machinery; it establishes the machine as the final arbiter of human competence. When mechanized calculation becomes the supreme criterion of reason, human judgment is demoted to erratic variance. The imitation game flips on its head: after building machines to simulate human behavior, institutions demand that human beings conform to the speed, metrics, and disciplines of the machine.

The imitation game remains a curious historical benchmark. It shows what an observer can deduce from written text alone. But surviving that encounter tells us nothing about who owns the system, what the system knows, or who bears the cost when it fails. Exposing the betrayal keeps the focus where it belongs: on the political decisions made once the game is over.

151, no. 2 (2022): 272-287.

- Gonçalves, Bernardo. "The Turing Test is a Thought Experiment." *Minds and Machines* 33, no. 1 (2023): 1-32.
- Levesque, Hector J. "On Our Best Behaviour." *Artificial Intelligence* 212 (2014): 27-35.
- Searle, John R. "Minds, Brains, and Programs." *Behavioral and Brain Sciences* 3, no. 3 (1980): 417-424.
- Sterrett, Susan G. "Turing's Two Tests for Intelligence." *Minds and Machines* 10, no. 4 (2000): 541-559.
- Turing, Alan M. "Computing Machinery and Intelligence." *Mind* 59, no. 236 (1950): 433-460.
- Vonnegut, Kurt. *Player Piano*. New York: Charles Scribner's Sons, 1952.

5 References

- Brynjolfsson, Erik. "The Turing Trap: The Promise & Peril of Human-Like Artificial Intelligence." *Daedalus*