Do public attitudes matter?

Extended abstract

Blay Whitby 07/03/2026

Declaration: no AI was used in the preparation of this abstract

Significantly, in his 1950 publication [1], Turing never said that he was offering some sort of operational test of whether or not a machine could think. Unfortunately for 75 years, various versions of such a misreading of Turing have had an unhelpful distracting influence on Artificial Intelligence (AI) research. The highly mistaken notion that deceiving humans by superficially imitating them in a short conversation is a test of thinking, or of intelligence, or even of consciousness, has undoubtedly held back development in AI.[2] Too much research effort has been expended on developing better chatbots, rather than pursuing useful AI or increasing the scientific understanding of intelligence. This misdirection has undoubtedly caused neglect of many more scientifically and philosophically important questions about thinking machines.

A rather more informed reading of Computing Machinery and Intelligence might be that success in the imitation game is a necessary but not sufficient indicator of a machine's ability to think. Rather than being some sort of operational test, a fairly small amount of success in the imitation game becomes nothing more than an illustrative example of the sort of performance that we might reasonably expect of a machine capable of thinking. This reading of Turing would have been rather more useful to the development of AI than the mistaken view that superficially simulating human conversation *of itself* constitutes thinking.

That the tendency to interpret success in the imitation game alone as an indicator of intelligence in machines is unhelpful to the development of AI has already been observed: including by the current author [3]. Deceit of humans in conversation is, at best, a very small part of the performance we might reasonably expect from a machine that is capable of thinking. If AI research focusses on the imitation game, much – probably most – of what should be examined in a scientific study of intelligence will be overlooked. Current developments in LLMs serve both to illustrate and to emphasize the consequences of this misreading of Turing.

In particular, machines designed solely to perform well in the imitation game neglect such aspects of thinking as homeostasis, the ability to handle a noisy environment, and the contribution of affective relations with that environment. This is far from an exhaustive list. Indeed, the reason it is hard to produce anything approaching a comprehensive list is that the above usual misreading of Turing has given the impression that simply *simulating* one very small aspect of thinking in a manner sufficiently accurate to *deceive* a human interrogator can somehow replace actually investigating the details of what thinking entails.

Even if we avoid the mistaken reading of Computing Machinery and Intelligence as describing an operational test for thinking machines, there remains Turing's claim that by 50 years after publication, when machines have become capable of deceiving an interrogator 70% of the time for 5 minutes, that public attitudes will have changed sufficiently that the

claim that machines can think will be generally accepted. This raises the obvious question "Should we assess the ability of machines to think by consulting public attitudes to the question?" The widespread use of LLM-based chatbots and the public acceptance of this technology make this question more urgent and significant.

Some very good reasons for not trusting public attitudes have emerged since Turing's landmark paper. In 1966, the very first chatbot – Weizenbaum's ELIZA [4] –introduced the notion of 'the ELIZA effect'. The ELIZA effect is the unavoidable tendency of human observers to interpret even vaguely human-like output as indicative of the ability to think. It has also been established that this is an unavoidable tendency. Even when people know that they are dealing with a machine they cannot help but relate to it in a human-like manner [5]. Perhaps the idea of trusting in the judgement of the public, as suggested by Turing, was always fundamentally flawed.

A further reason is the demonstration provided by the recent development of modern Large Language Models (LLMs) They amply illustrate that by the application of very large amounts of computational power and very large training sets that more and more convincing imitation of human-like output can be produced. If the future of AI is to be characterised only by the use of more computer power (with significant environmental costs) and larger training sets in order to simulate human conversation much will be missed. In particular research aimed at developing scientific understanding of the nature of intelligence – in humans, animals, and machines will be sidelined.

We should not perhaps, totally ignore public attitudes but neither should we take them as the main arbiter of scientific questions – particularly when we know that the ELIZA effect and similar psychological tendencies makes humans very poor judges *on this particular question*. There are good reasons for concluding that on the question of whether or not an artefact is capable of thought public attitudes are at best misleading. This is an area which deserves more scientific research, not more public-pleasing chatbots. Is there still life in the Turing Test? – very probably but there really shouldn't be.

<u>References</u>

[1] Computing Machinery and Intelligence, Turing, A. (1950), Mind 59, 433-60

[2] An example of such an interpretation of Turing 1950 can be found in Gunkel, D., (2012), The Machine Question, MIT Press p. 56:
"If a machine, Turing hypothesized, becomes capable of successfully simulating a human being in communicative exchanges with a human interlocutor, then that machine would need to be considered "intelligent". Although initially introduced for and limited to deciding the question of machine intelligence, the test has been extended to the question concerning personhood. "

[3] Hayes, P. and Ford, K. (1995) Turing test considered harmful
Proceedings of the 14th international joint conference on Artificial intelligence - Volume 1 p. 972 - 977

Whitby, B., (1996) The Turing Test: AI's Biggest Blind Alley? In Millican, P. and Clark, A. (eds) 1996 Machines and Thought: The Legacy of Alan Turing, OUP p. 53-62 repr in Whitby 1996, Reflections on Artificial Intelligence, Intellect Press p. 35-43

[4] Weizenbaum, J. (1966) ELIZA- A Computer Program for the Study of Natural Language Communication Between Man and Machine, Communications of the Association for Computing Machinery, 9 p. 36-45

[5] Hofstadter, Douglas R. (1996). The Ineradicable Eliza Effect and Its Dangers, Fluid Concepts and Creative Analogies: Computer Models of the Fundamental Mechanisms of Thought. Basic Books. p. 157.

Declaration of interest I was a participant in *not an organiser of* the 2012 Bletchley Park Turing Test, referenced in the cfp.