What is this?

This info-mural is one of seven "argumentation maps" in a series that explores Turing's question: "Can computers think and/or will they ever be able to?" Argumentation mapping is a method that provides:

- a method for portraying major philosophical, political, and pragmatic debates
- a summary of an ongoing, major philosophical debate of the 20th century
- a new way of doing intellectual history.

What does it contain?

Altogether the seven maps:

- summarize over 800 major moves in the debates threaded into claims, rebuttals, and counterrebuttals
- 97-130 arguments and rebuttals per map
- 70 issue areas in the 7 maps
- 32 sidebars history and further background

The argumentation maps:

- arrange debate so that the current stopping point of each debate thread is easily seen
- identify original arguments by over 380 protagonists worldwide over 40 years
- make the current frontier of debate easily identifiable
- provide summaries of eleven major philosophical camps of the protagonists (or schools of thought).

How do I get a printed copy?

You can order artist/researcher signed copies of all seven maps from www.macrovu.com for $500.00 plus shipping and handling.

② Can The Turing Test Determine Whether Computers Can Think?

The History and Status of the Debate — Map 2 of 7

An Issue Map™ Publication

1998

Start Here

1 Alan Turing, 1950
Yes, machines can (or will be able to) think. A computational system can possess all important elements of human thinking or understanding.

I believe that at the end of the century ... one will be able to speak of machines thinking without expecting to be contradicted.

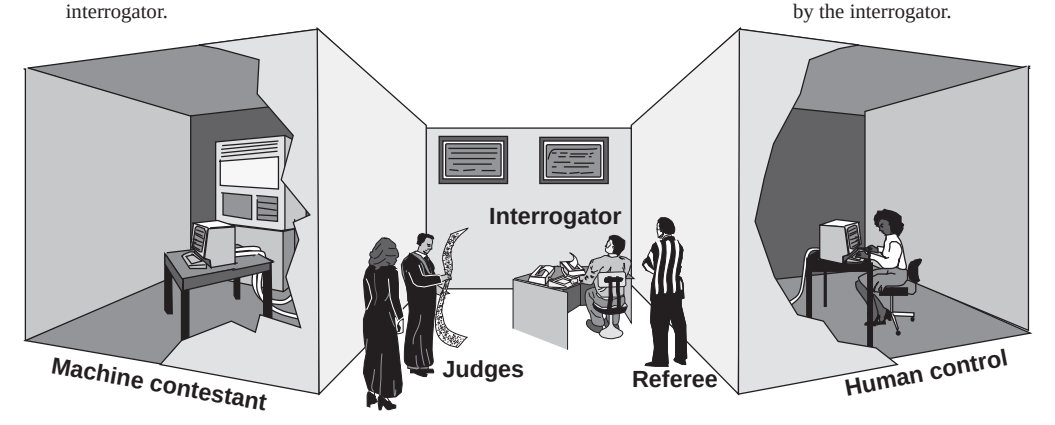

Alan Turing

2 Alan Turing, 1950
The Turing test is an adequate test of thinking. The Turing test is a test to determine whether a machine can think. If a computer can persuade judges that it is human via teletyped conversation, then it passes the test and is deemed able to think.
Notes:
- Shown here is a standard interpretation of Turing's test; many others are possible. Turing's original version of the Turing test takes the form of an imitation game (see "The Imitation Game," Box 4).
- Authors differ in whether they take the Turing test to establish "thinking" or "intelligence."

The Turing Test

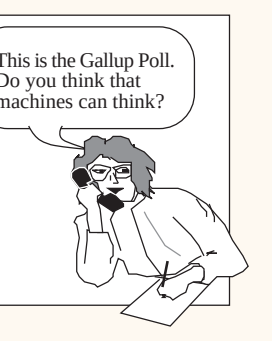
1 **In one room ...** a machine answers questions posed by the interrogator.

2 **In another room ...** a human control answers question posed by the interrogator.

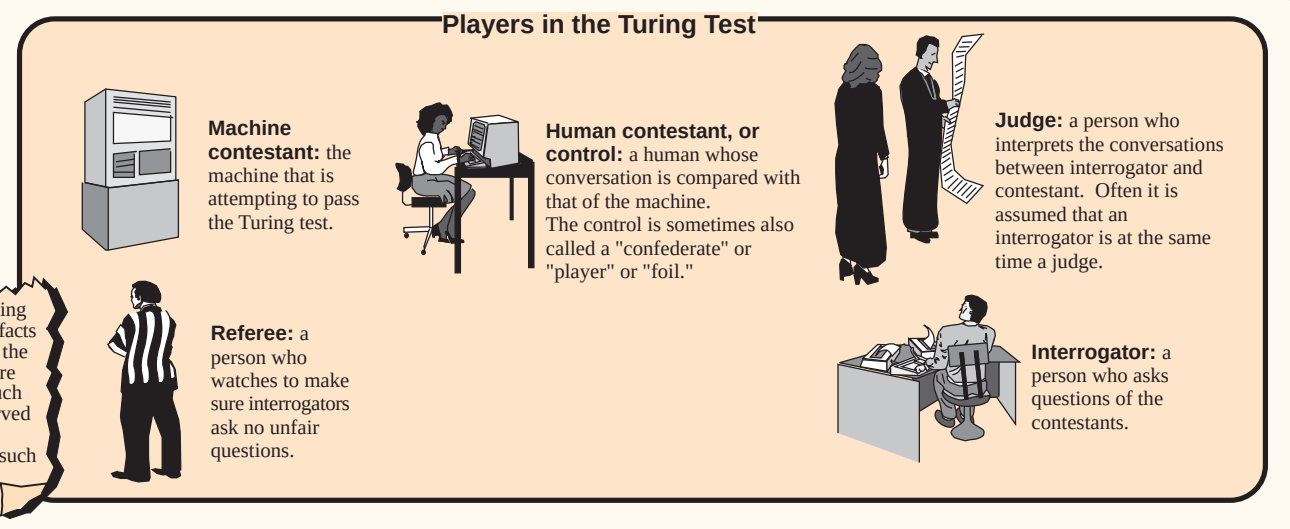


3 **In a third room ...** the interrogator engages in teletyped (computer) conversation with the contestants as judges look on. If a machine can trick the interrogator and the judges into thinking it is a human, then that machine has passed the Turing test.

3 Alan Turing, 1950
The question "can machines think?" is too vague. The question of whether machines can think is too vague. It cannot be decided by asking what these words commonly mean, because that reduces the question of machine intelligence to a statistical survey like the Gallup Poll. Hence, the question "can machines think?" should be replaced by the more precise question of whether they can pass the Turing test.



con · trol: A comparison case for what is being tested, to ensure that test results are not just artifacts of the test procedure or some other factor. In the Turing test the responses of human controls are compared with those of machines. Without such controls we would never know whether observed conversational ability was a result of actual conversational ability or of some other factor, such as the interrogator's bias, gullibility, etc.


Players in the Turing Test

Machine contestant: the machine that is attempting to pass the Turing test.

Human contestant, or control: a human whose conversation is compared with that of the machine. The control is sometimes also called a "confederate" or "player" or "foil."

Judge: a person who interprets the conversations between interrogator and contestant. Often it is assumed that an interrogator is at the same time a judge.

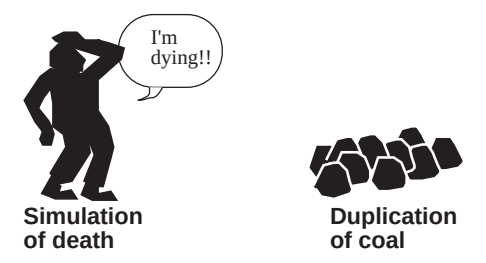
Referee: a person who watches to make sure interrogators ask no unfair questions.

Interrogator: a person who asks questions of the contestants.

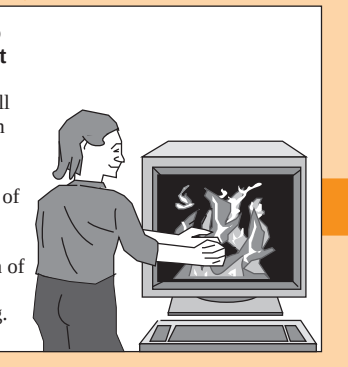
If a simulated intelligence passes, is it intelligent?

21 *Anticipated by* Jack Copeland, 1993
Simulated intelligence is not real intelligence. Just as simulated diamonds are not real diamonds, a machine that simulates intelligence to pass the Turing test is not really intelligent. In general, a simulated *X* is not a real *X*.

22 Jack Copeland, 1993
Some simulations are duplications. Two kinds of simulation can be distinguished. (1) Some lack essential features of the object simulated; for example, theatrical death lacks essential physiological features of real death. Such simulations are not duplications. (2) A second type captures the essential features of the object simulated despite being produced in a nonstandard way, for example artificial coal. Simulations of this kind are duplications.



23 John Searle, 1980b
Simulations are not duplications. A simulation of a fire will not burn. A simulation of metabolism in digestion will not nourish. A simulation of an automobile engine will not get anywhere. Similarly, a simulation of understanding is not actually understanding.



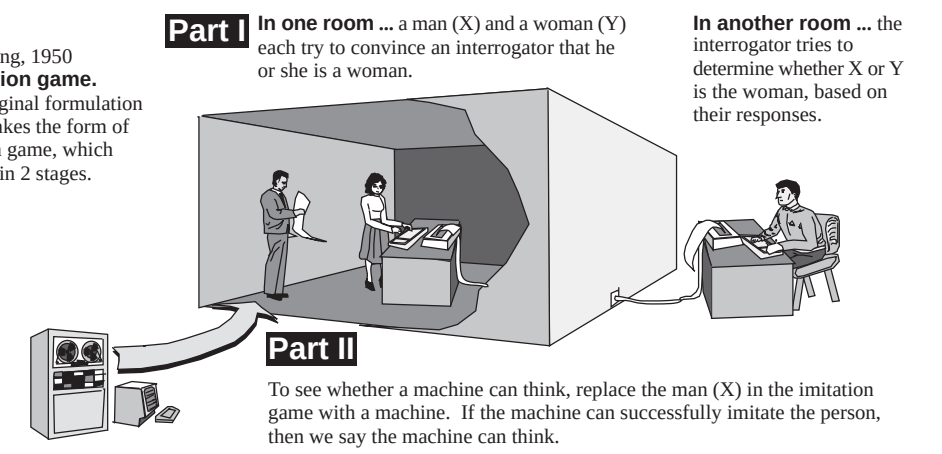
24 Lawrence Carleton, 1984
Simulations are duplications if the inputs and outputs are of the right kind. If a simulation uses the same kinds of inputs and outputs as the simulated phenomenon uses, then it is also a duplication. The Chinese Room, for example, uses the same kinds of inputs and outputs (i.e., symbol strings) to model speaking as humans use when they speak, hence it is a duplication as well as a simulation. Computer simulations of fire, digestion, etc., are different. Because they don't use the right kinds of input and output, they are mere simulations.
Supported by
"The Chinese Room Is More Than A Simulation," Box 55, Map 4.

Does the imitation game determine whether computers can think?

4 Alan Turing, 1950
The imitation game. Turing's original formulation of his test takes the form of an imitation game, which takes place in 2 stages.

Part I **In one room ...** a man (X) and a woman (Y) each try to convince an interrogator that he or she is a woman.

In another room ... the interrogator tries to determine whether X or Y is the woman, based on their responses.



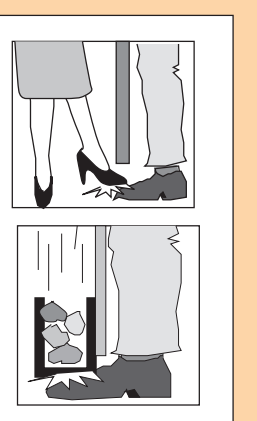
Interrogator: Will *X* please tell me the length of his or her hair?
X: My hair is shingled and the longest strands are about 9 inches long.
Y: I am the woman, don't listen to him!

Part II To see whether a machine can think, replace the man (X) in the imitation game with a machine. If the machine can successfully imitate the person, then we say the machine can think.

5 Kenneth Colby, F. Hilf, S. Weber, and H. Kraemer, 1972
The imitation game is flawed. The imitation game is a weak test for several reasons.
- The concept of "womanlikeness" is too vague.
- Control subjects may be inexperienced at deception.
- The test fails to specify what to do when a computer fails to imitate a man imitating a woman. Is the computer a successful imitation of a man failing to imitate a woman? Or is the computer a poor imitation of man? Or what?

6 Jack Copeland, 1993
A man simulating a woman is not a woman. If a man were to win the imitation game, he would only have succeeded in *simulating* a woman. He would obviously not *be* a woman. Therefore, the imitation game is not an adequate test.

7 Keith Gunderson, 1964
A box of rocks could pass the toe-stepping game. There is an analogy between the imitation game and a toe-stepping game, in which an interrogator puts his foot through an opening in the wall and tries to determine whether a man or a woman has stepped on his foot. There is a version of the toe-stepping game in which a box of rocks hooked up to an electric eye could win the toe-stepping game by producing the same effect as a human foot stepping on an interrogator's toe. However, a box of rocks cannot duplicate real human toe-stepping. Similarly, the imitation game just shows that a system can produce the outward effects of conversation, not that it can duplicate real human intelligence.

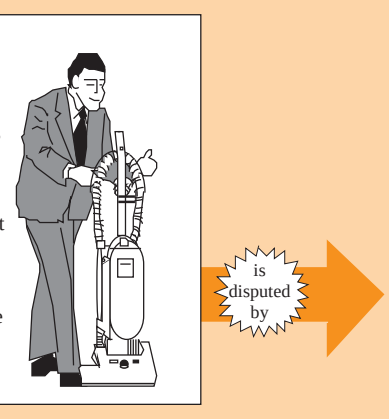


8 Keith Gunderson, 1964
One example cannot explain thinking. The imitation game attempts to explain thinking in terms of one example: a man's ability to imitate a woman. But this is just one of many different examples of mental activity, such as wondering, reflecting, deliberating, writing sonnets, and so forth. Thinking is "all purpose," whereas playing the imitation game uses just one ability.


Keith Gunderson

9 John Stevenson, 1976
The imitation game is an all-purpose test. Playing the imitation game is a "second order" ability that presupposes many other abilities. Hence, it is misleading to characterize the ability to pass the imitation game as just one example or property of thinking.

10 Keith Gunderson, 1964
The vacuum salesman example. If a vacuum-cleaner salesman demonstrated an all-purpose Swish 600 by sucking up dust from a carpet, the customer would be unsatisfied. He or she would want to see it suck up mud, straw, paper, and cat hairs from couches, tight corners, and so forth. One example is not enough to demonstrate the all-purpose Swish 600, just as one example is not enough to demonstrate thinking.

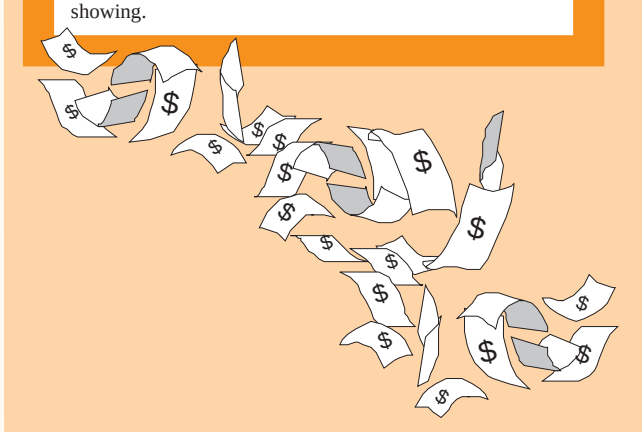


11 James Moor, 1987
The customer could infer that the vacuum was all purpose. The buyer would not have to actually watch the vacuum cleaner do everything in its repertoire to infer that it was a good vacuum. Even if the salesman only demonstrated the vacuum cleaner picking up dust, if the buyer saw that it came with various attachments, he or she could reasonably infer that they worked.
Supported by
"The Inductive Evidence Interpretation," Box 108.

Can the Loebner Prize stimulate the study of intelligence?

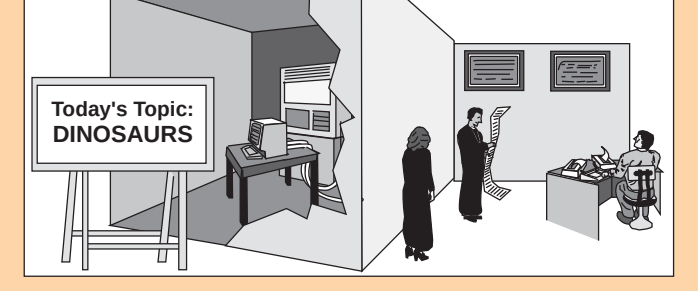
12
The Loebner Prize. A $100,000 award, funded by philanthropist Hugh Loebner, is offered for the first software to pass the Turing test in an annual contest started in 1990. Whether a system has passed the Turing test is determined in a competition where computers face off with humans in a special version of the Turing test. Even if entrants fail to pass the Turing test and win the $100,000 prize, an annual award is offered for the program that makes the best showing.

Hugh Loebner



13
The Loebner version of the Turing test. This version of the Turing test adds several measures to make it practical.
- Test questions are restricted to predetermined topics (see "Restricted Turing Tests," Box 14).
- Referees watch to make sure judges stick to the topics and ask fair questions without "trickery or guile."
- Prizes are awarded using a novel scoring method. Judges rank the contestants relative to one another, rather than just passing or failing them.
- The highest scoring computer in each contest is the yearly winner, and if a computer's score is equal to or greater than the average score of a human contestant, then it passes this version of the Turing test and wins the $100,000 prize.

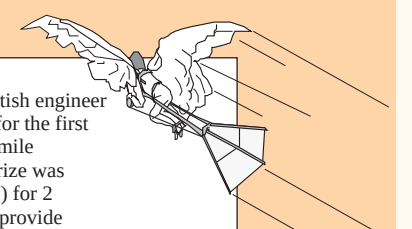
14
Restricted Turing tests. A kind of Turing test that restricts conversation to one topic, such as "women's clothes" or "automotives." Because restricted Turing tests are much easier to pass, they provide a feasible short-term goal for AI researchers to work toward.



15 Stuart Shieber, 1994a
Loebner's version of the Turing test lacks a clear goal. By restricting conversation to limited topics and using referees to ensure that no unfair questions are asked, the "open-ended, free-wheeling" nature of real human conversation is lost. As a result, contestants win by exploiting crafty tricks rather than by simulating intelligence; hence, the goal of understanding human intelligence is lost.

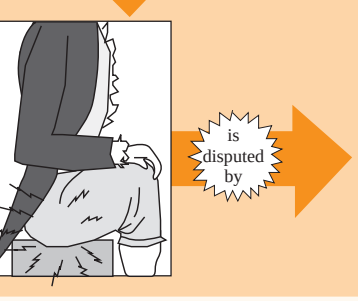
16 Stuart Shieber, 1994a
The Loebner Prize is too far beyond the reach of current technology. Because full-scale simulated conversation is beyond the reach of current technology, contestants in the Loebner competition are forced to exploit trickery to win the annual prize. The AI community should not focus on winning Loebner's "parlor game" version of the Turing test, but should engage in basic research that will incrementally lead the way toward artificially intelligent behavior.

17 Stuart Shieber, 1994a
The Kremer Prize. In 1959, British engineer Henry Kremer established a prize for the first human-powered flight over a half-mile figure-eight course. The Kremer Prize was successful (it was awarded in 1977) for 2 reasons. (1) It had a clear goal: to provide incentive to advance the field of human-powered flight. (2) The task was just beyond the reach of current technology: all the essentials of human-powered flight were available in 1959, they just had to be creatively combined.

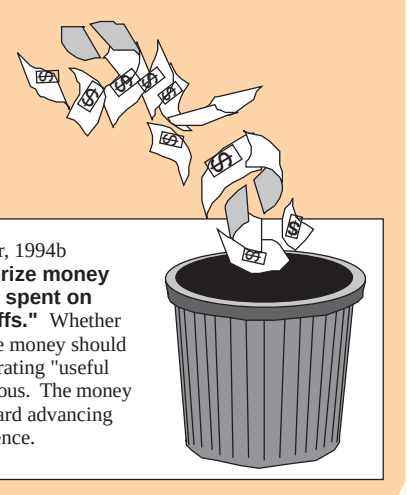


18 Stuart Shieber, 1994a
The da Vinci Prize. An imaginary prize is set up in 1442 (da Vinci's era) for the highest human-powered flight. We imagine that, like the Loebner Prize, a yearly competition was held and a prize was offered each year. At the time the prize was instituted, the requisite technology for human flight (in particular, the airfoil) was beyond the reach of current technology. As a result, competitors won by attaching springs to their shoes and jumping. Twenty-five years later no basic developments in human-powered flight had been achieved, but the trickery of spring-assisted flight had been highly optimized.

19 Hugh Loebner, 1994
Mozart's backside. After a coach ride to Vienna in 1781, Mozart complained that his buttocks were so tired that he could hardly ride on. Had a da Vinci Prize advanced the state of spring technology in earlier centuries, Mozart's coach might have boasted a smoother suspension, and his backside might not have been so sore. The Loebner Prize might turn out to be useful in unexpected ways, much as the da Vinci Prize might have.



20 Stuart Shieber, 1994b
The Loebner prize money should not be spent on "useful spin-offs." Whether the Loebner prize money should be spent on generating "useful spin-offs" is dubious. The money is best spent toward advancing artificial intelligence.



Legend

The arguments on these maps are organized by links that carry a range of meanings:

is supported by	Arguments that uphold or defend another claim. Examples include: supporting evidence, further argumentation, thought experiments, extensions or qualifications, and implemented models.
is disputed by	A charge made against another claim. Examples include: logical negations, counterexamples, attacks on an argument's emphasis, potential dangers an argument might raise, thought experiments, and implemented models.
is interpreted as	A distinctive reconfiguration of an earlier claim.
Unmapped Territory Additional arguments	This icon indicates areas of argument that lie on or near the boundaries of the central issue areas mapped on these maps. It marks regions of potential interest for future mapmakers and explorers.

Focus Box: The lowest-numbered box in each issue area is an introductory focus box. The focus box introduces and summarizes the core dispute of each issue area, sometimes as an assumption and sometimes as a general claim with no particular author.

Arguments with No Authors: Arguments that are not attributable to a particular source (e.g., general philosophical positions, broad concepts, common tests in artificial intelligence) are listed with no accompanying author.

Citations: Complete bibliographic citations can be found in the booklet that accompanies this map.

Methodology: A further discussion of argumentation analysis methodology can be found in the booklet that accompanies this map.

Anticipated by	Where this phrase appears in a box, it identifies a potential attack on a previous argument that is raised by the author so that it can be disputed.
As articulated by	Where this phrase appears in a box, it identifies a reformulation of another author's argument. The reformulation is different enough from the original author's wording to warrant the use of the tag. This phrase is also used when the original argument is impossible to locate other than in its articulation by a later author (e.g., word of mouth), or to denote a general philosophical position that is given a special articulation by a particular author.

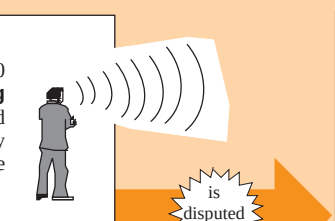
Version 1.0

The Issue Mapping™ series is published by MacroVU Press, a division of MacroVU, Inc. MacroVU is a registered trademark of MacroVU, Inc. Issue Map and Issue Mapping are trademarks of Robert E. Horn.

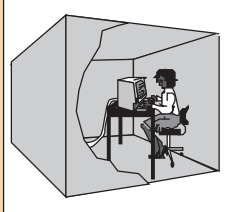
One of 7 in this Issue Mapping™ series—Get the rest!
The remaining 6 maps in this Issue Mapping™ series can be ordered with MasterCard, VISA, check, or money order. Order from MacroVU, Inc. by phone (206–780–9612), by fax (206–842–0296), or through the mail (Box 366, 321 High School Rd. NE, Bainbridge Island, WA 98110).

Other Turing test arguments

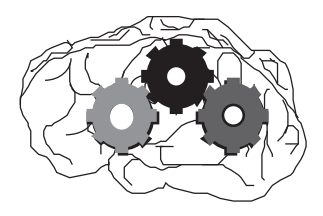
25 *Anticipated by* Alan Turing, 1950
ESP would confound the Turing test. Extrasensory perception could invalidate the Turing test in a variety of ways. If a competitor had ESP she could "listen in" on the judges and gain an unfair advantage. If a judge had ESP he could easily discern human from machine by clairvoyance. As Turing put it, "With ESP anything may happen."



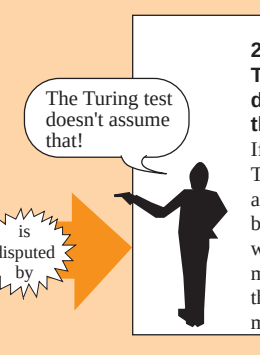
26 Alan Turing, 1950
Perform the test in telepathy-proof rooms. To prevent ESP-related problems, the test would have to be carried out in "telepathy-proof" rooms.



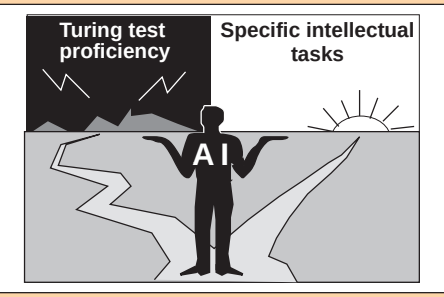
27 Michael Apter, 1971
Turing assumes that the brain is a machine. When Turing claims that machines can be taught in much the same way that children are, he assumes the brain is a machine. But whether or not the brain is a machine is part of what's at issue, because if brains were machines then obviously some machines (brains) would be able to think. So, Turing begs the question of whether machines can think—he assumes part of what he's trying to prove.



28 James Moor, 1976
The Turing test does not depend on the assumption that the brain is a machine. If a machine were to pass the Turing test, it wouldn't tell us anything about the relation between brains and machines. It would only show that some machines are capable of thinking, whether those machines are like brains or not.



29 B. Meltzer, 1971
The Turing test misleads AI research. The Turing test motivates AI researchers to accomplish a misleading goal—the general but arbitrary form of intelligence involved in carrying on a normal conversation. This kind of intelligence is arbitrary in that it is relative to particular cultures. AI researchers should focus their efforts on specific intellectual tasks that can be aggregated to form more generally intelligent systems, such as the ability to recognize patterns, solve differential equations, etc.



30 Bennie Shanon, 1989
The Turing test assumes a representationalist theory of mind. By confining itself to teletyped conversations, the Turing test assumes a representational theory of mind, according to which thinking involves computational operations on symbolic representations. But critics have pointed out various phenomena that cannot be accounted for in this framework. Hence, the Turing test begs the question of whether machines can think because its very procedure assumes that a representational system like a computer is capable of thinking. (Also see the "Can symbolic representations account for human thinking?" arguments on Map 3.)

31 Justin Lieber, 1989
Turing was not committed to representationalism. Turing leaves open the question of whether the representational account of intelligence is valid. In fact, in some passages he seems to reject the representational view in favor of connectionism (where he discusses learning machines). Moreover, Turing would have denied that an "autonomous" account of intelligence could be abstracted from the underlying mechanisms of intelligence, for "he was a hopeful mechanist through and through."

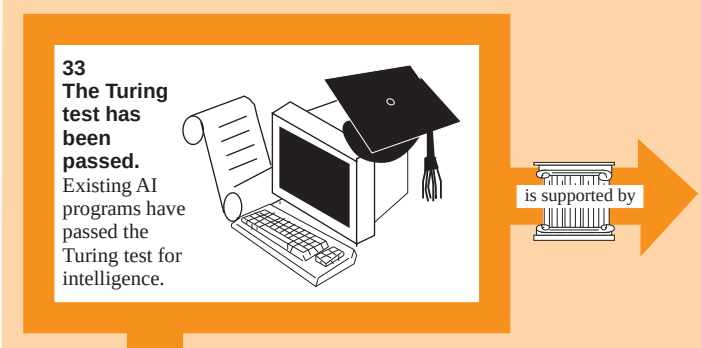
32 Bennie Shanon, 1989
The Turing test makes an autonomy claim. In assuming a representational theory of mind, the Turing test endorses an autonomy claim to the effect that intelligence can be studied solely in terms of representations and computations, without recourse to the brain, the body, or the world. But there is good reason to think the autonomy claim is false, and that in studying the mind we should also study its total embodied context.

au · ton · o · my (of theories): Two theories *A* and *B* are autonomous if the phenomena of *A* can be accounted for without recourse to the conceptual framework of *B*.

A: Phonology
B: Syntax
Everything in *A* can be explained without reference to *B* and vice versa.

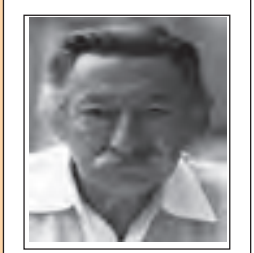
Have any machines passed the test?

33
The Turing test has been passed. Existing AI programs have passed the Turing test for intelligence.



Implemented Model

34 Joseph Weizenbaum, 1976
ELIZA. Written circa 1965, ELIZA emulates a Rogerian psychotherapist by using a few simple strategies. Although written in just 200 lines of BASIC code, ELIZA managed to fool many people, including professional psychotherapists, into believing it was intelligent. (See "The ELIZA Effect," Map 1, Box 106.)


Joseph Weizenbaum

Patient: Well, my boyfriend made me come here.
ELIZA: Your boyfriend made you come here.
Patient: He says I'm depressed much of the time.
ELIZA: I am sorry to hear that you are depressed.

Implemented Model

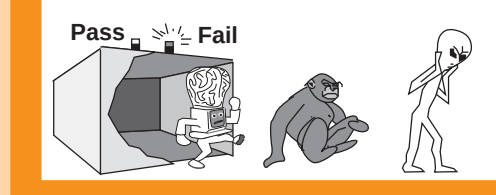
35 Kenneth Colby, 1981
PARRY. PARRY is a system that passed an extended version of the Turing test. PARRY simulates a paranoid subject by using a natural language parser and an "interpretation-action" module that tracks a range of emotional states and applies a set of paranoia-specific strategies for dealing with shame-induced distress. PARRY was tested against control subjects suffering from clinical paranoia, and the transcripts were judged by professional clinical psychologists.

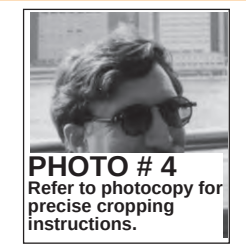
PARRY: ... I think he is behind the gangsters who are out to get me.
Interrogator: Did you see things that others don't seem to see?
PARRY: No, I don't. I'm not like the other patient here.
Interrogator: You didn't answer my question.
PARRY: Maybe it wasn't worth answering.

37 Robert Abelson, 1968
The extended Turing test. An extended version of the Turing test facilitates its use for scientific purposes. Its innovations are:
- The judges are not informed that a computer is involved in the test at all. They are told simply to compare the performance of 2 entrants with respect to some dimension, for example womanliness, paranoid behavior, etc.
- The computer does not participate on its own, but (unbeknownst to the judges) shares time with a human subject. The 2 are switched off periodically throughout the test.
- The judges score the human/computer pair after every question, so that separate scores can be generated for each. The point of this method is to force judges to focus on the specific dimension being evaluated.

Is failing the test decisive?

36 Ned Block, 1981
Failing the test is not decisive. It is possible to fail the Turing test for intelligence and still be an intelligent being.

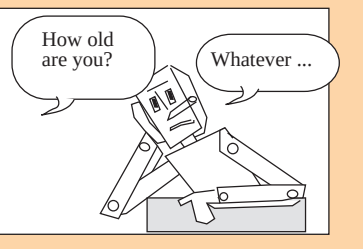




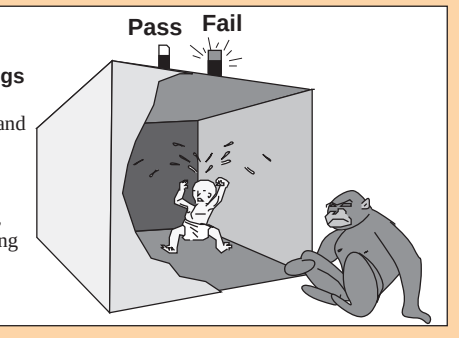
Ned Block

38 Ned Block, 1981
Judges may discriminate too well. Overly discerning or chauvinistic judges might fail intelligent machines solely because of their machinelike behavior.

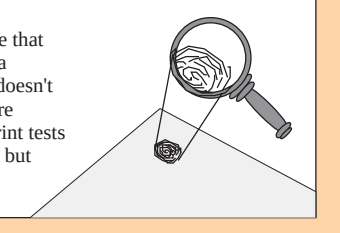
39 Jack Copeland, 1993
Intelligent machines could fail the test. An intelligent machine could fail the Turing test by acting nonhuman, by being psychologically unsophisticated (though intelligent), or simply by being bored with the proceedings.



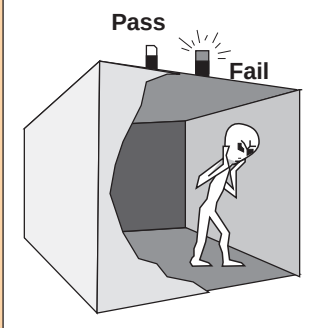
40 Jack Copeland, 1993
Some intelligent beings would fail the test. Chimpanzees, dolphins, and prelinguistic infants can think but would fail the Turing test. If a thinking animal could fail the test, then presumably a thinking machine could too.



41 Jack Copeland, 1993
Nondecisive tests can be useful. It is true that failing the Turing test is not decisive—it is not a definitive litmus test for intelligence. But that doesn't make it a bad test. Many perfectly good tests are nondefinitive but useful. For example, fingerprint tests are nondefinitive (they can be foiled by gloves) but remain useful nevertheless.



42 Robert French, 1990
The Turing test is species biased. Failing the Turing test proves nothing about general intelligence because only humans can pass it. Intelligent nonhuman entities (such as intelligent machines, aliens, etc.) could never pass the test, because they would not be able to answer subcognitive questions. Such questions can only be correctly answered on the basis of a specifically human associative network of representations.

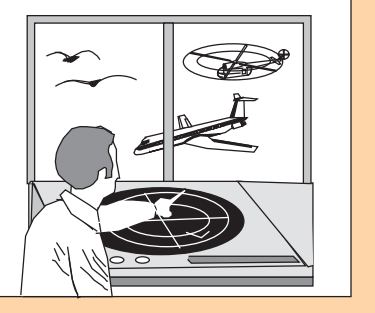



Robert French

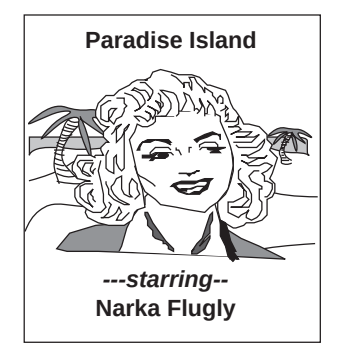
43 *Anticipated by* Robert French, 1990
We can make the test fair by excluding subcognitive questions. A machine's inability to answer subcognitive questions (e.g., questions about ranking associative pairings) may show that machines can never match human responses, but we can make the test unbiased by excluding such questions.

44 Robert French, 1990
A purely cognitive Turing test is not an intelligence test. We cannot distill a level of purely cognitive questions to make an unbiased test. To exclude all subcognitive questions would be to exclude all questions involving analogy and categorization, which would render the test useless as a test of what we call intelligence.

45 Robert French, 1990
The seagull test is species biased. The seagull test—an imaginary test to determine whether a machine can fly—shows how a Turing-type test is species biased. Judges compare 2 radar screens that track a seagull and a putative flying machine, respectively. A machine passes if the judges can't distinguish it from the seagull on the radar screens. But the seagull test tells us nothing about flying in general, because only seagulls (or machines behaving exactly like them) could pass. Non-seagull flyers (such as airplanes, helicopters, bats, etc.) could never pass the test, because they would produce a distinctive radar signature.

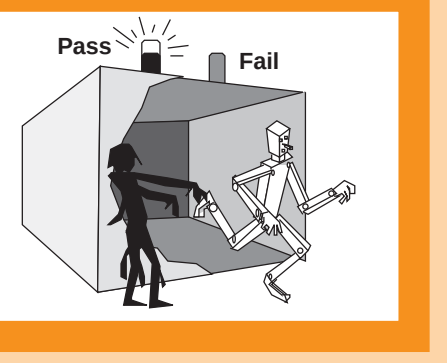


46 Robert French, 1990
Rating games reveal the subcognitive underpinnings of human intelligence. Rating games provide an experimental basis for formulating subcognitive questions that only humans could answer. A rating game asks a subject to judge the strength of an "associative pairing" between 2 terms. For example, in the "neologism rating game" a subject is presented with a new term, like "Flugly," and asked to rate it as a name for various things:
(a) a child's favorite teddy bear,
(b) the name of a glamorous movie star, or
(c) the name of character in a W. C. Fields movie.
Only humans could answer such questions plausibly, on the basis of subcognitive associations.

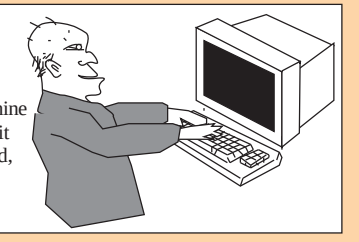


Is passing the test decisive?

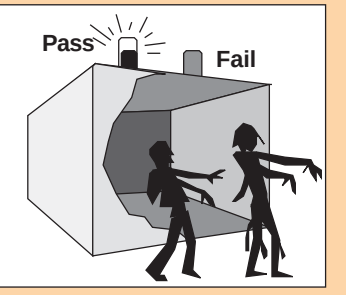
47 Ned Block, 1981
Passing the test is not decisive. Even if a computer were to pass the Turing test, this would not justify the conclusion that it was thinking intelligently.
Note: Because the issue of passing the Turing test is closely tied with the issue of what can be inferred from external behavior, this region contains arguments that are similar to those in the "Is the test, behaviorally or operationally construed, a legitimate intelligence test?" arguments on this map.



48 Charles Karelis, 1986
Nonstandard human controls can mislead a judge. An unintelligent machine could pass the Turing test if it were compared with a stupid, flippant, or in general anomalous control.

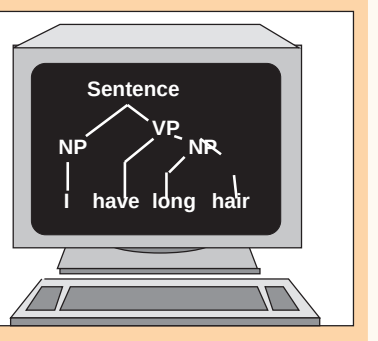


49 Charles Karelis, 1986
A zombie or an unconscious machine could pass the test. A behaviorally sophisticated but unconscious zombie could pass the Turing test, but it would not be thinking. Similarly, a behaviorally sophisticated but unconscious computer might pass the Turing test without thinking.

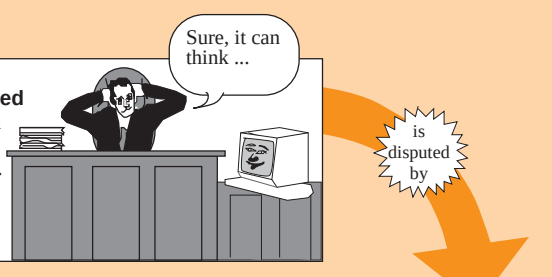


50 Charles Karelis, 1986
Consciousness is essential to thinking. No matter how complex an entity's behavior is, it's not thinking unless it's conscious. Proposed counter- examples (e.g., problem solving in a dreamless sleep) are just deviant uses of language.
Note: For more on consciousness, see Map 6.

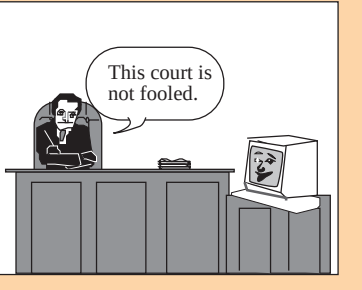
51 John Barresi, 1987
Merely syntactic machines could pass the Turing test. Purely syntactic machines could be developed that would pass the Turing test. But without a full naturalistic semantics it is doubtful that such machines will ever think in the human sense. A naturalistic semantics would provide a global model of situational context that could allow flexible interaction with the world as well as sustained self-perpetuation.
Supported by
"The Cyberiad Test," Box 68.



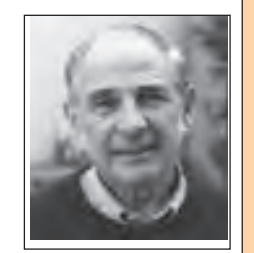
52 Ned Block, 1981
Human judges may be fooled too easily. Human judges can be fooled by unintelligent but cleverly programmed machines.
Supported by
"The ELIZA Effect," Map 1, Box 106.



53 James Moor, 1987
A serious judge would not be easily fooled. A critical judge would not lazily interact with a competitor in the Turing test, but would be focused on distinguishing the human from the computer. Such a judge could not easily be fooled.



54 John Searle, 1980b
The Chinese Room passes. The Chinese Room thought experiment involves a system that passes the Turing test by speaking Chinese but fails to understand Chinese.
Note: Also, see Map 4.

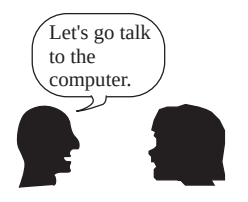

John Searle

55 Joseph Rychlak, 1991
The anthropomorphizing objection. Our natural tendency to anthropomorphize machines should caution us against calling them intelligent. We are easily fooled into thinking that a machine that engages in conversation can think (see "The ELIZA Effect," Map 1, Box 106). But, passing the Turing test isn't enough to establish intelligent thought. Introspection is also necessary, and the Turing test doesn't reveal it.

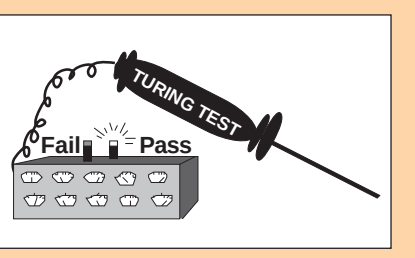

Joseph Rychlak

56 Joseph Rychlak, 1991
The reverse Turing test. Humans tend to treat machines as fellow humans, making quick and unwarranted ascriptions of intelligence. We say that computers "want to do" things or "will not allow" certain alternatives. This tendency could be tested for using a "reverse Turing Test." Instead of testing to see if a machine can trick a human into thinking it is intelligent, we test to see if a machine can avoid being treated as intelligent. We ask: "Can a machine interacting with a machine avoid being anthropomorphized?" (p. 60)
Note: Also, see "The ELIZA Effect," Map 1, Box 106.

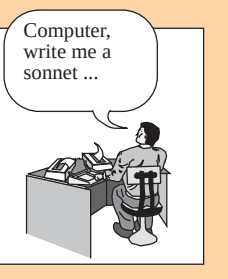
57
The test is too narrow. The Turing test is limited in scope. It narrowly focuses on one ability—the ability to engage in human conversation—and is therefore inadequate as a test of general intelligence.
Supported by
"One Example Cannot Explain Thinking," Box 8.



58 Daniel Dennett, 1985
The quick probe assumption. Passing the Turing test implies indefinitely many other abilities, for example the ability to understand jokes, write poetry, discuss world politics, etc. It is a "quick probe" of general intelligence.



59 James Moor, 1976, 1987
Narrowness objections play a misleading numbers game. It is misleading to characterize the Turing test as just one test, because the test can be used to conduct many different kinds of tests. The interrogator can test for the ability to tell jokes, speak a foreign language, discuss urban renewal, or for any of a wide range of activities.

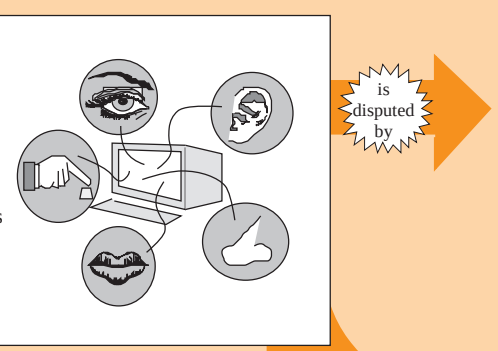


60 James Moor, 1987
Linguistic evidence is sufficient for good inductive inference. By considering only linguistic behavior, we can make an inductive inference that a system can think. Of course, we may revise this inference as further evidence comes to light, but that does not mean that all such evidence must be gathered before a justified inference can be made. Otherwise, "scientists would never gather enough evidence for any hypothesis" (p. 1128).
Supported by
"The Inductive Interpretation," Box 108.

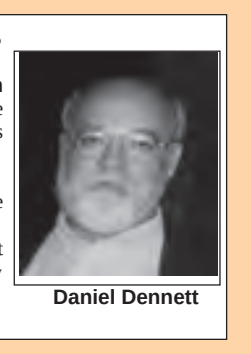

James Moor

61 Jerry Fodor, 1968
The Turing test only provides partial evidence of intelligence. Passing the Turing test does not guarantee successful simulation of human behavior. Any given Turing test provides only a finite sample of a test taker's behavioral repertoire. Therefore, passing the Turing test only provides partial, inductive evidence of intelligence.

62 Peter Carruthers, 1986
The sense organs objection. The Turing test neglects perceptual aspects of intelligence. A machine might pass the test by producing competent verbal behavior but fail to understand how its words relate to the perceptual world. An adequate test would have to provide the machine with some sort of sensory apparatus.

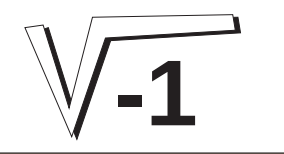


63 Daniel Dennett, 1985
Sense organs are an unnecessary addition to the test. Because the Turing test indirectly tests sense-based capacities (see "The Quick Probe Assumption," Box 58), the addition of sensory devices is unnecessary. It doesn't make the test any more demanding than it already is.

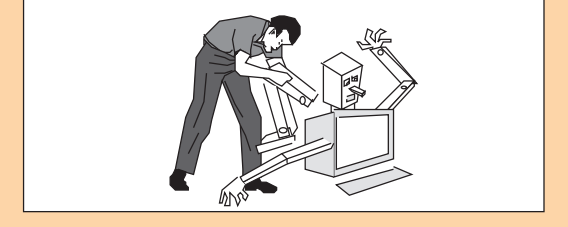

Daniel Dennett

64 Jack Copeland, 1993
Sense organs were not prohibited by Turing. The Turing test does not prohibit machines with sense organs from entering. In fact, Turing thought the best way to prepare a computer for competition would be to give it sensory abilities and an appropriate education.

65 Jack Copeland, 1993
Understanding can be tested without sensory interaction. Understanding can be tested through verbal quizzing, even when sense organs are lacking. Indeed, there are certain kinds of concepts that have to be tested for through verbal quizzing because of their nonperceptual nature (e.g., "square root," "intelligence," and "philosophy").

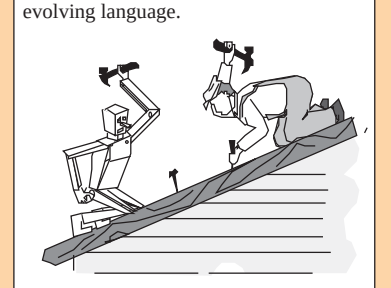


66 Steve Harnad, 1995
The Turing test underdetermines the goal of creating humanlike robots. By focusing on a particular subset of human behavior (symbolic capacity), the Turing test underconstrains the engineering problem of creating robots with human capabilities. Because of this, machines that pass the Turing test will not be easily extensible to machines that have other human capacities.



67 John Barresi, 1987
You can't fool Mother Nature. Because the Turing test only measures the ability of a machine to trick judges, it cannot test for the flexible adaptations needed to interact in the real world, where Mother Nature is the judge of intelligence. In the real world intelligence is best indicated by an ability to survive through coordinated perception and action. The Turing test neglects such abilities and biases us to think of intelligence in an intellectualist way.

68 John Barresi, 1987
The Cyberiad test. A race of machine people ("cybers") is judged intelligent if they can do whatever humans can do in a natural environment. They must be able to replace humans in their social roles, perpetuate their species for as long as the human species could, and maintain an evolving language.

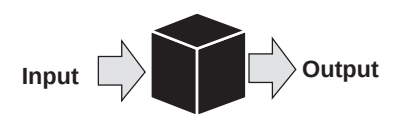


Is the test, behaviorally or operationally construed, a legitimate intelligence test?

69
The Turing test provides a behavioral/operational definition of intelligence. The Turing test defines thinking in terms of overt, measurable behavior. It offers a behavioral/operational definition of our ordinary concept of thinking.
Disputed by
"A Box of Rocks Could Pass the Toe-Stepping Game," Box 7.
Note: A similar debate takes place in the "Is passing the test decisive?" and "Is failing the test decisive?" arguments on this map, which deal with the question of what can be demonstrated by successful simulation.

Unmapped Territory
Additional behaviorism/ operationalism arguments

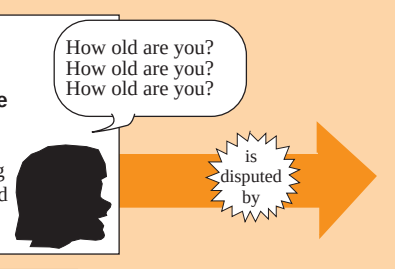
70 Jack Copeland, 1993
The black box objection. The Turing test treats the mind as a "black box." Internal processes are left hidden and mysterious; all that matters is input and output. But internal processing should not be ignored, because it is crucial to determining whether or not a system really thinks.




Jack Copeland

71 Jack Copeland, 1993
SUPERPARRY. An advanced alien civilization could write a program that, like Colby's PARRY program (see "PARRY," Box 35), would contain ready-made responses to questions. This SUPERPARRY would be better prepared for the test, however, because it would contain responses to all possible questions. SUPERPARRY could therefore pass the Turing test by brute force. But it is clearly not thinking; it is just dumbly searching a long list of conversations (see "The All-Possible-Conversations Machine," Box 90).

72 *Anticipated by* Jack Copeland, 1993
The repeated sentence maneuver. An interrogator could trick SUPERPARRY by typing the same sentence over and over again.

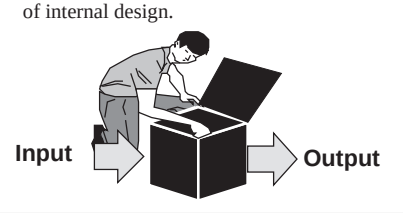


73 Jack Copeland, 1993
The machine is prepared for repeated sentences. The aliens would design the program to anticipate repeated sentences and to reply with "increasingly irritable responses," ultimately terminating altogether.

74 *Anticipated by* Jack Copeland, 1993
The substitution trick. An interrogator could trick SUPERPARRY by typing, "Let's write *!= in place of 'horse'; now tell me whether a *!= has a tail."

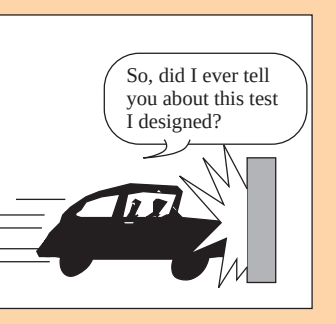
75 Jack Copeland, 1993
The machine is prepared for substitution tricks. Because there are only a finite number of "substitution tricks," SUPERPARRY could be built handle all of them.

76 Jack Copeland, 1993
The beefed-up Turing test. The addition of a design criterion to the original Turing test ensures that the system being tested has the right kind of internal design.



77 Alan Turing, 1950; James Moor, 1987
Knowledge of internal processes is unnecessary. Inferences about thinking are not based on knowledge of internal operations. We generally infer that someone thinks just on the basis of outward behavior. "Not even brain scientists examine the brains of their friends before attributing thinking to them" (Moor, 1987, p. 1129).

78 Jack Copeland, 1993
Turing in a car crash. Suppose you and Alan Turing are driving to a party and your car crashes. He is gashed up, but there is no blood. Instead, wires and silicon chips stick out of his skin, and you wonder if he is just an elaborate puppet. In such a case you might modify your judgment that Turing is a thinking being based on this new knowledge of his internal constitution. So, internal processes *are* relevant to judgments of thinking.



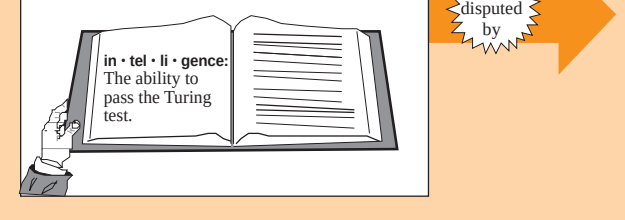
79 John Searle, 1980b
Overt behavior does not demonstrate understanding. The Chinese Room argument shows that a system can engage in Chinese-speaking behavior yet fail to understand Chinese. The Chinese Room is a system that could pass the Turing test without thinking in Chinese.
Note: Also, see Map 4.

be · hav · ior · ism: A school of psychology that takes the overt behavior of a system to be the only legitmate basis of research. We can distinguish philosophical (or logical) behaviorism from methodological behaviorism. Philosophical behaviorism is the position that mental states are reducible to sets of input-output correlations (or "behavioral dispositions"). Methodological behaviorism is a research program that restricts itself to the investigation of behavior and its observable causes.

80 James Moor, 1976, 1987
The behavioral/operational interpretation is vulnerable to counterexamples. Whether behaviorally or operationally interpreted, the Turing test is vulnerable to cases where unthinking machines pass the test or thinking machines fail it (see the "Is passing the test decisive?" and "Is failing the test decisive?" arguments on this map). But, if the Turing test is interpreted as a source of inductive evidence, then such counterexamples can be accommodated.
Supported by
"The Inductive Evidence Interpretation," Box 108.

op · er · a · tion · al · ism: A philosophical position asserting that scientific concepts should be defined in terms of repeatable operations. For example, intelligence can be defined in terms of the ability to pass the Turing test.

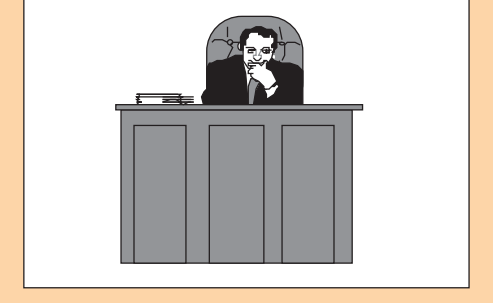
81 *Anticipated by* Ned Block, 1981
The operationalist interpretation. An intelligent system is operationally defined as one that passes the Turing test. By definition, systems that pass are intelligent; systems that fail are unintelligent.



82 Ned Block, 1981
The operational interpretation is too rigid. If thinking is operationally defined in terms of the ability to pass the Turing test, then systems that pass are necessarily intelligent and systems that fail are necessarily unintelligent. But, this is too rigid. For example, an intelligent machine could fail the test (see the "Is failing the test decisive?" arguments on this map) and unintelligent machines could pass the test (see the "Is passing the test decisive?" arguments on this map).
Supported by
"The Neo–Turing Test Conception of Intelligence," Box 88.

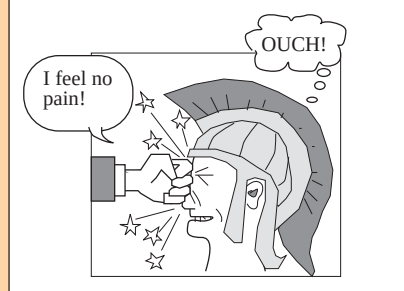
83 *Anticipated by* Ned Block, 1981
The behavioral disposition interpretation. A system is intelligent if it is behaviorally disposed to pass the Turing test. In this interpretation, neither passing the test nor failing the test is conclusive, because intelligence doesn't require passing but only a disposition to pass.

84 Ned Block, 1981
Human judges are unreliable. Even though the behavioral disposition argument avoids the problems associated with the operational interpretation, it still has the problem that it relies on human judges. Some judges may chauvinistically reject intelligent machines (see "Judges May Discriminate Too Well," Box 38), whereas others might be overly liberal in passing cleverly designed but unintelligent machines (see "Human Judges May Be Fooled Too Easily," Box 52).
Supported by
"The Neo–Turing Test Conception of Intelligence," Box 88.

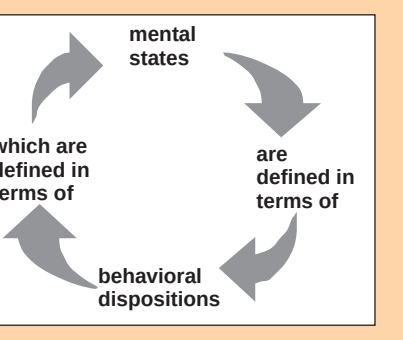


85 Gilbert Ryle, 1949
Philosophical (or logical) behaviorism. What we call intelligence is simply patterns of potential and actual behavior. The external behavior of a system is enough to determine whether or not it thinks.

86 Hilary Putnam, 1975a
Super Spartans. We can imagine a race of Super Spartans who have been trained to suppress any behavioral indication of pain. They exhibit no pain behavior whatsoever, even when they are in excruciating pain. So, philosophical behaviorism fails to give an account of what pain is.

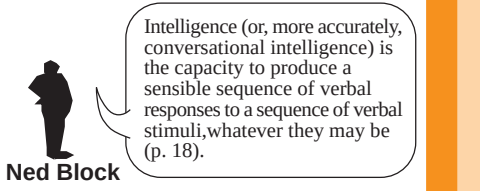


87 Roderick M. Chisholm, 1957
Philosophical behaviorism is circular. Any attempt to define mental states in terms of behavioral dispositions relies on further mental states. But, this is a circular definition of mentality. For example, if the desire to eat chicken is defined as a disposition to eat chicken when chicken is present, then this presupposes a belief that chicken is edible, which presupposes a belief that chicken is nonpoisonous, and so forth.



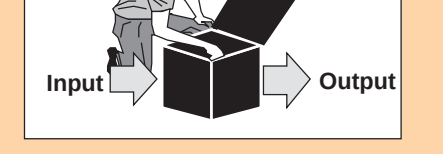
Is the neo–Turing test a legitimate intelligence test?

88 *Anticipated by* Ned Block, 1981
The neo–Turing test conception of intelligence. The neo–Turing test conception of intelligence is a beefed-up version of behaviorism that avoids classical problems associated with behaviorism as well as problems associated with the Turing test. It does not rely on human judges (it merely requires "a sensible sequence of verbal responses to verbal stimuli") and the test avoids problems associated with dispositional analysis by substituting "capacities" for dispositions.
Note: Block sets up this position in order to attack it, so that he can show that even a beefed-up version of behaviorism will fall prey to psychologism.

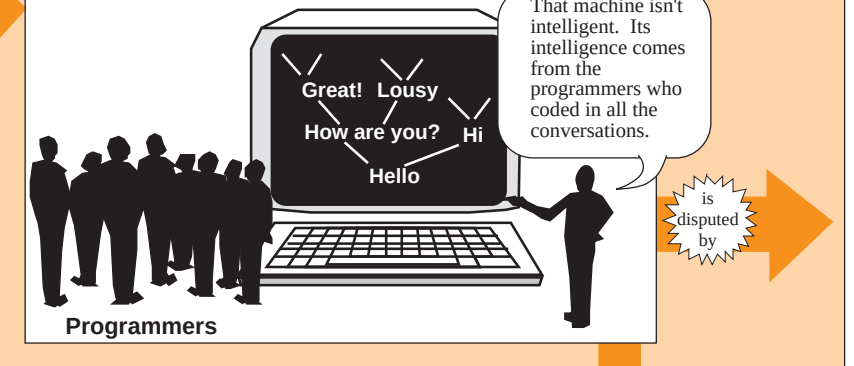


Ned Block

89 Ned Block, 1981
Psychologism: Internal differences matter. Psychologism is the doctrine that whether or not behavior is intelligent depends on the nature of the information processing that produces it. If 2 systems behave in exactly the same way, one might be intelligent while the other is stupid, because of differences in the way each system processes information. Behavior by itself isn't enough; internal differences matter.



90 Ned Block, 1981
The all-possible-conversations machine. A machine could be built to engage in sensible conversation by searching a database containing all possible lines of conversation in a finite Turing test. The machine would be intelligent according to the neo–Turing test, but because of the way it processes information, it is clearly unintelligent. It merely echoes the intelligence of its programmers. Note: Block simply calls this the "unintelligent machine."



91 Robert Richardson, 1982
The neo–Turing test fails to improve on behaviorism. Block's neo–Turing test is subject to the following dilemma.

Either	Or
Internal states are allowed as part of the definition of intelligence through the notion of a capacity, *in which case* the test is not purely behavioral as Block claims.	Internal states are not allowed to play a role in the definition of intelligence, *in which case* we have no reason to believe that Block's conception fares any better than classical behaviorism. (See "Super Spartans," Box 86; "Philosophical Behaviorism Is Circular," Box 87.)

In Either Case
The neo–Turing test fails to offer a plausible behavioral conception of intelligence.

be · hav · ior · al dis · po · si · tion: A specification of input-output correlations that can be used to define mental terms. Here is an example of a definition of hunger for chicken, stated in terms of behavioral dispositions.

If I see chicken ...
then I eat the chicken.

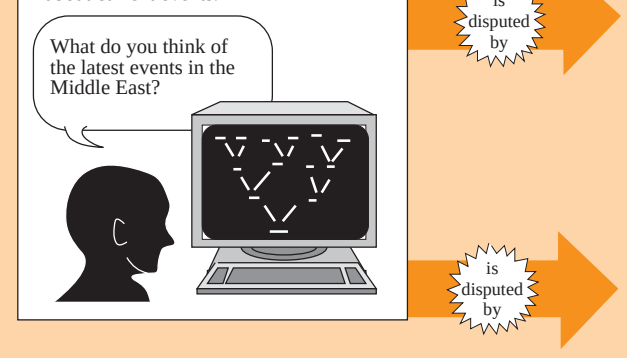
be · hav · ior · al ca · pac · i · ty: A specification of input-output correlations that includes a specification of internal states (beliefs, desires, etc.). Here is an example of a definition of hunger for chicken, stated in terms of behavioral capacities.

If I see chicken and I believe chicken is edible
and I desire chicken ...
then I eat the chicken.

92 *Anticipated by* Ned Block, 1981
All intelligent machines exhibit the intelligence of their designers. There is nothing unusual about the all-possible-conversations machine. It may be said of any intelligent machine that its intelligence is really just that of its designers.

93 Ned Block, 1981
Some machines can exhibit their own intelligence. The all possible-conversations machine, which dumbly searches a list of conversations, only exhibits the intelligence of its programmers. A machine equipped with general mechanisms for learning, problem solving, etc. would exhibit its own intelligence as well as the intelligence of its designers.

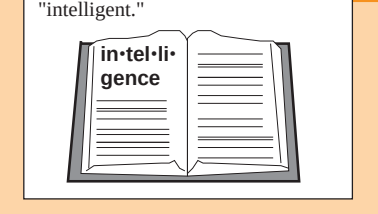
94 *Anticipated by* Ned Block, 1981
The machine is dated. The machine would not be able to answer questions about current events.



95 Ned Block, 1981
A dated system can still be intelligent. Intelligence does not require knowledge of current events. For example, the all possible-conversations machine could be programmed to simulate Robinson Crusoe, who is intelligent even though he can't answer questions about recent events.

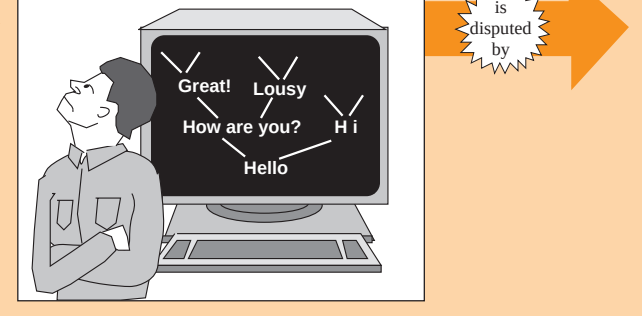
96 Ned Block, 1981
The machine could be updated. Programmers could periodically update the machine's list of responses.

97 *Anticipated by* Ned Block, 1981
The all-possible-conversations machine redefines intelligence. We normally conceive of intelligence relative to input-output capacity, not relative to internal processing. By stipulating an internal processing condition on intelligence, Block merely redefines the word "intelligent."



98 Ned Block, 1981
The machine is all echoes. Intelligence is not being redefined, because it is part of our normal conception of intelligence that input-output capacity can be misleading. Someone who plays grand-master chess by copying another grand master would not be deemed a grand master. Similarly, the all-possible-conversations machine, which merely echoes its programmers, would not normally be considered intelligent.

99 *Anticipated by* Ned Block, 1981
Assuming we are better than the machine is chauvinistic. A system that processes information differently than we do may not be intelligent in our sense, but we may not be intelligent in its sense either. To assume that human intelligence is better than machine intelligence is chauvinistic.

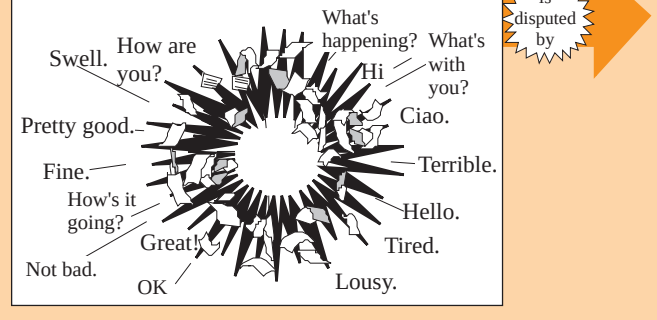


100 Ned Block, 1981
Intelligence requires "richness" of information processing. The all-possible-conversations machine isn't unintelligent simply because it processes information differently than we do. Its style of information processing is what's at issue. In particular, it lacks the richness of processing capacity that is associated with intelligence.

101 *Anticipated by* Ned Block, 1981
What if we are like the all-possible-conversations machine? Suppose humans process information in the same way that the all-possible-conversations machine does. Would that mean humans are unintelligent?

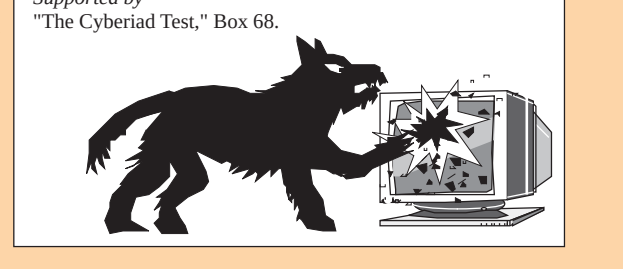
102 Ned Block, 1981
The machine would be intelligent. Admitting that the all-possible-conversations machine would be intelligent if it were like us is compatible with claiming that it is not in fact intelligent. It's like asking an atheist what he would say if he turned out to be God. He might admit that if he turned out to be God, God would exist, but deny that God in fact exists.

103 *Anticipated by* Ned Block, 1981
Combinatorial explosion makes the all-possible-conversations machine impossible. Programming the machine to engage in an hour-long Turing test is impossible, because it would give rise to a "combinatorial explosion." The programmer would have to code more strings than there are particles in the universe.



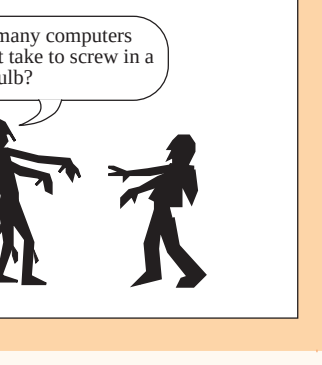
104 Ned Block, 1981
The all-possible-conversations machine is logically possible. The all-possible-conversations machine is logically possible even if the machine is empirically impossible. Because the neo–Turing test makes a claim about the *concept* of intelligence, it is refuted by the logical possibility of the all-possible-conversations machine.

105 John Barresi, 1987
The all-possible-conversations machine cannot pass the Cyberiad test. The Cyberiad test avoids the problems posed by Block's all-possible-conversations machine because the Cyberiad test tests for more than just linguistic ability. The Cyberiad test tests for survival in a natural environment, and hence linguistic competence is not enough to pass it.
Supported by
"The Cyberiad Test," Box 68.



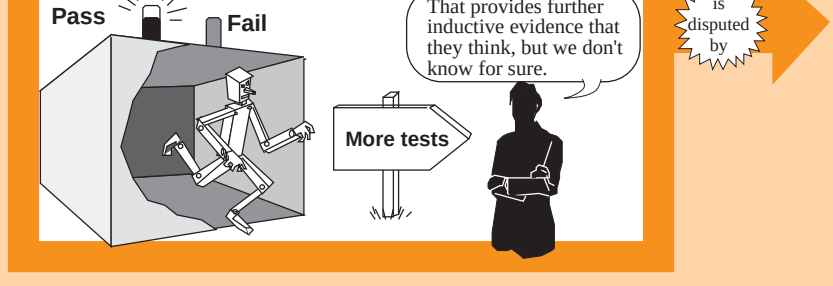
106 Robert Richardson, 1982
A brute list searcher cannot flexibly respond to changing contexts. The all-possible-conversations machine could not adapt its speech to novel and changing contexts. A machine that searches a list to produce responses to verbal stimuli might be able to find some response that would make sense in some contexts, but it could not coherently adapt to the contingencies of an ongoing conversation.

107 Charles Karelis, 1986
Bantering zombies can't think. Bantering zombies, using processing mechanisms far superior to those of the all-possible conversations machine, would still be unintelligent. No matter how much we enrich the information processing of the all-possible-conversations machine, it still can't think unless it's conscious.
Supported by
"Consciousness is Essential to Thinking," Box 50.



Is the test, as a source of inductive evidence, a legitimate intelligence test?

108 James Moor, 1976, 1987
The inductive evidence interpretation. The Turing test provides inductive evidence that a system thinks. Such evidence may be weak or strong, but it is never deductively certain. The claim that a system thinks, based on its passing the Turing test, remains subject to revision as further evidence comes to light.



109 Douglas Stalker, 1978
The inductive interpretation is really an explanatory interpretation. The inductive evidence interpretation assumes that a computer's passing the Turing test should be explained in terms of its thinking. But, thinking is not the best explanation of a computer's ability to pass the test. A better explanation is a mechanical one, in terms of the structure of the computer, its program, and its physical environment.

110 James Moor, 1978
Competing explanations can be compatible. An explanation of a computer's behavior in terms of its thinking may be compatible with an explanation in terms of its underlying mechanisms. A computer that prints payroll checks can be explained both in terms of its physical mechanisms and its abstract program—these are compatible explanations of the payroll computer's behavior. Similarly, mental and mechanical explanations of a computer that passes the Turing test may be compatible.