


Rapaport W (2006), Turing Test. In: Keith Brown, (Editor-in-Chief) *Encyclopedia of Language & Linguistics, Second Edition*, volume 13, pp. 151-159. Oxford: Elsevier.

Between Vico and Herder (*see* **Herder, Johann Gottfried (1744–1803)**), announcing the German creativity of sign (Tetens's *Bildendes Dichtungsvermögen*), Turgot invented and spontaneously articulated several features of a change leading to romanticism.

*See also:* Condillac, Etienne Bonnot de (1714–1780); Diderot, Denis (1713–1784); Herder, Johann Gottfried (1744–1803); Rask, Rasmus Kristian (1787–1832).

## Bibliography

Droixhe D (1979). 'Un plan inédit de Turgot pour un discours sur l'origine, la formation et le mélange des langues.' *Marche romane 29*, 207–222.

Droixhe D (1987). 'Turgot: L'économie du langage poétique.' In *De l'origine du langage aux langues du monde*. Tübingen: Narr.

Droixhe D (1988). 'Turgot, commercio e filiazione.' In Formigari L & Lo Piparo F (eds.) *Prospettive di storia della linguistica*. Rome: Ed. Riuniti.

Droixhe D (1989). 'Le primitivisme linguistique de Turgot.' In *Primitivisme et mythe des origines*. Paris: Presses de l'Université de Paris-Sorbonne.

Hassler G (1976). 'Ansatze zur Diskussion um ein sprachliches Relativitätsprinzip in der Auseinandersetzung Turgots mit Maupertuis.' *ZPhon. 29*, 491–494.

Hassler G (1985). 'Parallelismen zum ökonomischen Denken in der Sprachtheorie Turgots.' *Wiss. Zeitschr. Univ. Halle 34*, 80–87.

Piron M (1961). *Etymologie*. Brugge: De Tempel.

Rosiello L (1987). 'Turgot's *Etymologie* and modern linguistics.' In Buzzetti D & Ferriani M (eds.) *Speculative grammar, universal grammar and philosophical analysis of language*. Amsterdam: Benjamins.

Swiggers P (1990). 'Le fondement cognitif et sémantique de l'étymologie chez Turgot.' *Cahiers Ferdinand de Saussure 43*, 79–89.

# Turing Test

**W Rapaport**, State University of New York at Buffalo, Buffalo, NY, USA



## Introduction

'The Turing Test' is the name given to any of several variations of an experiment proposed by the mathematician and computer scientist Alan M. Turing in his essay, 'Computing machinery and intelligence,' which appeared in the philosophy journal *Mind* in 1950. The Test, briefly put, is for an interrogator to decide whether a participant in a natural-language conversation is a human or a computer. The participant passes the test to the extent that it convinces the interrogator that it is human (even if it is really a computer). (For simplicity, I will usually call the one (or two) participant(s) in a Turing Test other than the interrogator *the participant*. In different versions of the Test, the participant(s) might be a male and/or female human, a computer, or perhaps some other candidate for the ascription of *intelligence*.)

The experiment, a version of a parlor game that Turing called *the imitation game*, was designed to help answer the question 'Can machines think?' by replacing that informal question (containing the vague terms *machines* and *think*) by a more precise question: Can a computer 'pass' a Turing Test? Thus, the vague terms *machines* and *think* are replaced by, respectively, the more precise notions of a digital computer (or a Turing machine) and the more precise – or at least behaviorally determinable – notion of passing the Test. (This strategy – replacing a vague, informal question that is difficult to answer with a more precise question that is easier to answer – is similar to the strategy that Turing (1936) adopted in giving his extremely influential answer to the question of what 'computable' means.)

### The Imitation Game

The original 'imitation game' consists of three players: a man, a woman, and an interrogator, each in a separate room, communicating only via teletype (or what today would be considered an 'instant messaging' interface), such that none of the players can see or hear the others, thus limiting their interactions to typed (i.e., written) linguistic communication. We will consider the extent to which this is a limitation later on. The woman's task is to convince the interrogator that she is the woman; the man's task is to **imitate** a woman so as to convince the interrogator that he **is** the woman; and the interrogator's task is to determine who is which. The man wins the imitation game (i.e., 'passes' the imitation-game 'test') if he succeeds in convincing the interrogator (e.g., perhaps by deception) that he is the woman; thus, the woman succeeds (as does the interrogator) to the extent that

the man fails. This version of the imitation game seems to be computationally solvable: a computational study has suggested that it is possible to identify stylistically whether the author of a given text is male or female (see Argamon *et al.*, 2003) for a computational investigation of this.

### The Turing Test and Its Variations

Several versions of the Turing Test may be found in Turing's essay and in subsequent literature. Turing's first version merely replaces the man in the imitation game by a computer. If the interrogator fails as often as he or she did with the imitation game, then the computer is said to have 'passed the Turing Test'. But does this mean that the computer passes the Test if it is able to convince the interrogator that it is a man who, in turn, is trying to convince the interrogator that he is the woman?

It is reasonably clear from Turing's essay that he had in mind something less convoluted. Thus, the second version of the Test replaces the man by a computer and the woman by another **man**, and so the computer's task is to convince the interrogator that it is a man (see Colby *et al.*, 1972: 202 and Shieber, 2004: 100–104 for further discussion of the gender issue).

A more gender-neutral version replaces the man by a computer and the woman by a **human** (of either sex). The computer passes this version of the Test if it convinces the interrogator that it is the human.

In modern versions of the Test, this is simplified even more, by having an interrogator and only one participant: either a computer or a human. Again, the computer 'passes' this Test if it convinces the interrogator that he or she is interacting with a human. Turing (1950) alludes to this version as a *viva voce*, i.e., an oral examination; a subject passes an oral exam by convincing the examiner that he or she 'understands' the material that he or she is being tested on.

Another variation (Stuart C. Shapiro, personal communication) depends on whether the interrogator **knows** that there is a possibility that the participant is a computer or whether the interrogator is merely being asked to judge whether the participant is intelligent (or as intelligent as a normal person).

Turing (1950: 441) also mentions a version of the Test in which the interrogator must distinguish between a digital computer and a discrete-state machine. Other variations are described later on.

### Passing the Test

As a standard for deciding whether a participant has passed the Test, Turing (1950: 442) proposed that "an average interrogator will not have more than 70 per cent. [*sic*] chance of making the right identification after five minutes of questioning." Some researchers have complained that this sets the standard too low, but the actual statistical standard to be applied could easily be varied (e.g., 70% increased to 95%, 5 minutes increased to several months, etc.). Turing's prediction was "that at the end of the century [i.e., by 2000], the use of words and general educated opinion will have altered so much that one will be able to speak of machines thinking without expecting to be contradicted" (Turing, 1950: 442). (We will examine these two standards later in this essay. At the end of the century, in fact, several anthologies considered the status of Turing's prediction: Akman and Blackburn, 2000; Moor, 2000–2001; Moor, 2003; Shieber, 2004.)

The Turing Test has been likened to two, more familiar situations:

1. The two-entity version (interrogator and participant) is like the situation in which a human corresponds with a pen pal (Harnad, 1989). In real life, the human would be unlikely to doubt that the pen pal was a human. One point of the Turing test is that if the pen pal were, in fact, a computer, still the human would not question its humanity (i.e., its human level of intelligence or its human ability to think).
2. As Stuart Shieber (2004) points out, the interrogator's task is like that of someone determining whether something is 1 m long by comparing it with a standard meter; both items share the same length. Here, the interrogator is determining whether something has a certain feature (namely, a human level of intelligence) by seeing if it shares that feature with (i.e., if it can be measured against) another entity, which, by hypothesis, is a 'standard' for having that feature. And, in the case of the Turing Test, the standard is written linguistic communication (Shieber's term is 'verbal behavior').

## Experiments with Real Turing-like Tests

Two early AI (artificial intelligence) programs in the areas of natural-language processing and cognitive modeling are often mentioned in connection with the Turing Test. Joseph Weizenbaum's (1966, 1967) computer program Eliza appears to converse in more or less fluent English with its user, apparently simulating a Rogerian psycho-therapist. Many anecdotes of people being fooled into thinking that they were conversing with a real human suggest that Eliza has passed a Turing-like Test. Yet Eliza's natural-language-processing techniques are mere pattern

matching using little or no grammatical knowledge (e.g., swapping pronouns such as you and me in certain patterns) – i.e., the techniques are not cognitive. Hence, Weizenbaum has claimed that Eliza is a counterexample to the Turing Test. But this has not been subject to controlled experiments.

On the other hand, Kenneth Mark Colby's computer program Parry is an Eliza-inspired program that, using more sophisticated natural-language-processing techniques than Eliza, simulates a paranoid patient. Parry has been the subject of controlled Turing-Test-like experiments (Colby *et al.,* 1972; Colby, 1981), and Colby, at least, claims that it has passed the Turing Test, having convinced human psychotherapist judges that they were conversing with a human paranoid patient. On the other hand, that is all that Parry can discuss, and so it is not clear that the sort of test that it passes is a full-fledged Turing Test.

The Loebner Prize Competition, held annually since 1991, is a Turing-like test that has offered a monetary award "for the first computer whose responses were indistinguishable from a human's. Each year an annual prize of $2000 and a bronze medal is awarded to the most human computer. The winner of the annual contest is the best entry relative to other entries that year, irrespective of how good it is in an absolute sense" (Loebner, 2003). Most of the winners have been elaborate Eliza-like programs (sometimes called *chatterbots*). However, one notable exception was David Levy's CONVERSE program (Batacharia *et al.,* 1999), which won the Loebner Prize in 1997, and which was developed by a team including at least one well-known researcher in computational linguistics (Yorick Wilks). The Loebner Competition has been critically analyzed by the computational linguist Stuart Shieber, who has argued that the Competition, unlike other competitions for professional computational linguists (e.g., the Message Understanding Conferences and the Text Retrieval Conferences), has not fostered research and development, but merely encouraged unintelligent, Eliza-like chatterbot programs whose sole goal is to fool the judges (Shieber, 1994a, 1994b; for Loebner's reply, see Loebner, 1994).

Turning the tables, some projects have used computers as 'interrogators' in a Turing-like test to unmask computer programs that attempt to convince other computer programs that they are human and hence entitled, say, to e-mail accounts. However, these tests are not based on natural-language understanding or even verbal behavior; they usually involve visual images containing distorted words that only humans can read (Ahn *et al.,* 2004). One formal investigation that replaces the interrogator with a Turing machine claims to show that no computer can be a perfect interrogator in the standard Turing Test (Sato and Ikegami, 2004). And there have been other investigations of the formal, logical, and computational structure of the Turing Test (e.g., Bradford and Wollowski, 1994; Shieber, 2005).

## What Must a Computer Be Able to Do in Order to Pass the Turing Test?

A significant feature of the Test is the central role of language. To pass the Test, a participant must be able to communicate in the written natural language of the interrogator to such an extent that the interrogator cannot distinguish its linguistic behavior from that of a human using that language. (To simplify matters, we will assume that all participants speak that language equally well.)

This feature echoes a historical antecedent to the Turing Test: in his *Discourse on method* (1637: Part V), the philosopher René Descartes argued that there was a test that could show that a machine that looked and behaved like a human was not a real human, namely, that even if it could use language in certain appropriate, responsive ways (e.g., to "ask what we wish to say to it" if "touched in a particular part," or to "exclaim that it is being hurt" if touched somewhere else), it could not "arrange its speech . . . in order to reply appropriately to everything that may be said in its presence" (Descartes, 1637/1970: 116). There are two issues to consider: (1) Is linguistic ability a reasonable replacement for the informal notions of 'thinking' or 'intelligence'? (2) Must the test be limited to linguistic ability?

### Is Written Linguistic Behavior a Reasonable Replacement for 'Cognition'?

'Thinking' is a vague term, and 'intelligence' is not necessarily, or even probably, intended in the sense of the IQ-test notion of intelligence, but merely as a convenient synonym for 'cognition'. Therefore, we will use the latter term, which is less controversial.

To approach an answer to the question of whether written linguistic behavior can replace 'cognition', we can look at what Turing thought the computer might have to be able to do. In his examples of the computer's behavior and in his replies to a series of potential objections, Turing suggests that the computer might have to do the following (kinds of) things:

- Answer questions about its physical appearance and abilities.
- Answer questions about its personality and experiences.
- Write and discuss a poem on a given topic.

- Do arithmetic.
- Solve chess problems.
- Have access to, and be able to use, commonsense information about the seasons, Christmas, familiar fictional characters, etc. (This is sometimes called world knowledge.)
- Learn.

For Turing, the last of these was the most important. Presumably, this learning would have to come from reading, not from doing, although Turing also suggests that the machine might have other "experience[s], not to be described as education." In any case, Turing seems to have been willing to restrict the Test-passing capabilities of the machine to "all purely intellectual fields," which, presumably, manifest themselves through language.

Insofar as ordinary cognition is surely involved in these sorts of activities, being able to do them does seem to be a reasonable test of human-level cognition. Surely, if a machine could do such things, an interrogator would be justified in being unable to distinguish between it and a human who could do them equally well. Hence (perhaps), the interrogator would be justified in concluding that the participant had cognition. Turing, however, does not say that the interrogator needs to draw the conclusion that the participant really has cognition (i.e., really can do these things); perhaps this is why he wrote that "the use of words and general educated opinion will have altered so much that one will be able to speak of machines thinking without expecting to be contradicted."

Thus, a computer must have at least the following two capacities to pass a Turing Test: (1) the ability to understand and generate natural language; (2) lots of world knowledge. Let us consider each of these.

**Understanding Natural Language** What counts as 'understanding' (including generating) natural language? Turing apparently thought that it simply meant behaving in such a way that the interrogator **believes** that the participant understands the language, i.e., that the interrogator **attributes** natural-language understanding to the participant (cf. Dennett's (1968) "intentional stance"). Cognitivists hold that it means something more than mere behavior. Perhaps any entity that exhibits such behavior must have the further capacity to be able to parse linguistic input, construct a (mental) representation of it, integrate or assimilate it into 'background knowledge', determine an appropriate linguistic response, and formulate and express that response in a grammatical utterance. Or **understanding** natural language might mean something even more than certain behavior together with this natural-language-**processing** capacity: Perhaps the entity must have "intrinsic intentionality" or consciousness (Searle, 1980).

**World Knowledge** Understanding natural language is probably not something that can be done merely on the basis of linguistic knowledge (e.g., knowledge of a grammar and lexicon). It probably requires much world knowledge – not only semantic and pragmatic knowledge, but also something like what the educational reformer E. D. Hirsch (1987) calls 'cultural literacy' or what the AI researcher Douglas B. Lenat (1995) calls 'commonsense knowledge', i.e., lots of facts about the world, and perhaps also some specialized domain knowledge (e.g., knowledge of how to play chess).

The computational linguist Terry Winograd argued for the claim that natural-language understanding requires such knowledge: consider these two sentences, which differ in only one word: (a) "The city councilors refused the demonstrators a permit because they advocated violence" and (b) "The city councilors refused the demonstrators a permit because they feared violence." Arguably, it is 'world knowledge' or 'commonsense knowledge' about the roles of city councilors and demonstrators, as well as semantic knowledge about the meanings of the words *advocated* and *feared*, that aids in determining that the word *they* refers to the demonstrators in (a) but to the city councilors in (b).

Moreover, someone reading *The Diary of Anne Frank* who does not understand the word *Gestapo* might be told that it was Hitler's secret police, but if the reader does not know who Hitler was, this will not be of much help. The reader needs to know that Hitler was the Nazi dictator of Germany during World War II, who the Nazis were, where Germany is, what World War II was, etc.

### Is Written Linguistic Communication a Limitation?

Must the test be limited to linguistic behavior? After all, cognition in general is usually taken to consist of more than just linguistic ability; e.g., it includes learning, perception, reasoning, memory, planning (and acting on those plans), etc. In the original imitation game, the restriction to linguistic interrogation eliminated the need to deal with the physical make-up of the participants. Limiting the interrogation to linguistic communication does not seem debilitating: surely the sorts of conversations that Turing suggested cover a wide range of intellectual activities – enough to convince an interrogator of the

participant's cognition (or, more precisely, the participant's cognitive-behavioral indistinguishability from a human).

Linguistic ability by itself is surely a mark of cognition. Moreover, it may be that linguistic ability is the essential core of the kind of cognition that is in question, in the sense that any entity that had such linguistic ability is highly likely to have – or require – most other kinds of cognition (Rapaport, 1988; Harnad, 1989; Rey, 2002). Not only does natural-language understanding require world knowledge, but it is arguably the case that natural-language understanding is 'AI-complete'; i.e., 'solving' the natural-language-understanding problem may require solving all other problems in AI (artificial intelligence). If so, then, insofar as passing the Turing Test requires natural-language understanding, it would be a sufficient condition of cognition. In any case, the psychologist Stevan Harnad (1989) has proposed a Total Turing Test variation that extends the Test to allow for demonstrations of perception, action, and other sorts of non-linguistic cognition (or cognitive indistinguishability from humans).

If passing the Turing Test requires, not merely the ability to fool or (statistically) convince an interrogator that he or she is conversing with a human (or at least with an entity with human-level cognition), but actual, human-level linguistic abilities, then what kinds of abilities are required? Research in AI and computational linguistics since the mid-1950s suggests that any natural-language understander – including a Turing-Test-passing program – must be able (at least) to:

1. take coherent discourse (as opposed to isolated sentences) as input
2. understand ungrammatical input
3. make inferences and revise beliefs
4. make plans (for speech acts, to ask and answer questions, and to initiate conversation)
5. understand plans (including the speech acts of all interlocutors in the conversation)
6. construct user models (i.e., 'theories of mind', or psychological theories, of all interlocutors)
7. learn about the world and about language, and do this in part via a natural-language interaction
8. have background (or 'world', or 'commonsense') knowledge
9. remember what it heard, learned, inferred, and revised.

Arguably, any entity having these abilities could be said to be able to understand language or, more generally, to be cognitive, or to think. The open research question is whether **computational** theories of these cognitive abilities can be developed.

## What Does Passing the Turing Test Mean?

### Does Passing the Turing Test Demonstrate Real Cognition or Only the Appearance of Cognition?

Is a machine that **appears** to understand natural language or to have cognition **really** understanding natural language; does it really have cognition? Perhaps the early-to-mid-20th-century behavioral bias made it seem reasonable to Turing that his test answers this question in the only way possible. But the modern late-20th-to-early-21st-century cognitive bias has kept the question open.

One position is that if the output of the machine is not produced by 'the same' methods that humans use, then although its external behavior might be similar to a human's, its internal mechanism is not and, hence, it is not really doing them. Another position is that anything that is capable of these complex sorts of linguistic behavior is indeed communicating in language (at least, no less than a human is) and not merely appearing to, no matter what internal mechanism produces the output. That is, it is not merely simulating the use of language; rather, at least in the case of language use (and other cognitive phenomena), this sort of simulation **is** the real thing (Rapaport, 1988).

Another line of argument, however, suggests that even in the case in which it is felt that the appearance is not reality, it is near enough so as not to matter. Some support for this is given by Turing's statement "that at the end of the century **the use of words** and **general educated opinion** will have altered so much that one will be able to speak of machines thinking without expecting to be contradicted." One way that this might come about is if "the use of words" were to change, e.g., the meaning of *think* might be extended metaphorically to cover both machines and humans. Enlarging the scope of meaning of a word is something that has happened many times before: "To fly" was once something that only birds and bats were said to do; now, airplanes, helicopters, space shuttles, and their human riders are all said to fly. A 'computer' was once a **human** whose job was to carry out complex or lengthy mathematical calculations; now, it is a **machine** that does what those humans once did. Such semantic enlargements can come about by metaphorical extension: perhaps planes do not literally fly (in the bird-sense), just as philosophers do not literally shoot down an opponent's argument, but we eventually come to apply the metaphor so often that it becomes an unquestioningly acceptable – even unavoidable – predication.

In addition to "the use of words" changing, it is also possible for "general educated opinion" about the nature of cognition to change, perhaps as a result

of progress in AI, philosophical analysis, or psychological research, e.g., the meaning of the word *think* might be abstracted to an essence that allows it to be predicated of both AI-programmed computers and humans (not to mention other animals). Our current theoretical understanding of heavier-than-air flight encompasses both bird flight and airplane flight. And our current theoretical understanding of computation – due, in fact, to Turing (1936) – has its foundation in an analysis of what human computers do. Extending this idea, progress in AI and cognitive science could cause us to have a better understanding of cognition in such a way that the new understanding applies to both humans and machines.

### Is Passing the Turing Test a Sufficient Demonstration of Cognition?

In what sense does passing the Turing Test **replace** the vague notion of 'thinking'? Logically, there are two possibilities: it is a necessary condition, or it is a sufficient condition (or both). To say that passing the Turing Test is a **necessary** condition of 'thinking' is to mean that if something thinks, then it can pass the Turing Test. It might, of course, be the case that something can think but fails the Test, perhaps because the interrogator is not up to par or because the thinking agent simply is having an 'off day'. But, presumably, under normal circumstances, a real cognitive agent that takes the Test will pass it.

More controversial is the proposition that passing the Turing Test is a **sufficient** condition of 'thinking'. To say this is to mean that if something passes the Test, then it 'thinks'. Several objections to the Turing Test take the form of showing that it is not a sufficient condition for thinking.

**The 'Aunt Bertha' Table-Lookup Objection** One, due to the philosopher Ned Block (1981), considers a computer program that stores a finite, but extremely large, number of possible conversations – all the one-hour conversations that the programmer's 'Aunt Bertha' might possibly have during her entire life, on any subject whatever. Such a program would, presumably, pass the Test, yet would not normally be considered to 'think'. The point is that the methods whereby the responses are produced are behaviorally, but not psychologically, plausible. Thus, we have a case of something that passes the Test but does not think, which is a counterexample to the test's sufficiency. Thinking, according to this objection, is not merely a matter of behavior, but also a matter of how that behavior is produced, and certain ways of producing it are not suitably cognitive (such as memorizing a large lookup table), whereas certain others might be (such as requiring reasoning; cf. Shannon and McCarthy, 1956).

One possible response is to deny the possibility of such a 'pre-compiled' lookup table (e.g., such a system could not learn new words). To this objection it can be replied that there is only a finite (albeit astronomically large) number of possible words, all of which could be pre-compiled. On behalf of the Test, another response to Block's objection is to admit this, but to point out that the Test still provides considerable evidence supportive of a claim to cognition and, in the absence of any ability to know how the cognitive behavior was produced (as in the pen pal case), that might be all the evidence one could possibly have.

**The Chinese-Room Objection** Perhaps the most well-known objection to the Test's sufficiency is the philosopher John Searle's (1980) Chinese-Room Argument. Here, a human who knows no Chinese is placed in a room with an instruction book that tells him how to manipulate certain symbols that are input to him in the room, and then how to output other symbols on the basis of these manipulations. The interrogator is a native speaker of Chinese who inputs questions written in Chinese characters and reads the output from the room. By hypothesis, the interrogator judges that the entity with which he is fluently conversing understands Chinese, whereas, again by hypothesis, the human in the room in fact knows no Chinese and is merely manipulating marks that are 'meaningless' to him. Thus, we seem to have a case in which something does not have a certain sort of cognition (in this case, understanding natural language) yet passes the Turing Test (and, moreover, via 'reasoning' as opposed to 'memorization').

One response to Searle's proposed counterexample is to object that the solo human is not the Test-taker; rather, the Test is being taken – and passed – by the system consisting of the human combined with the instruction book, and it is possible that that **system** does understand Chinese even though its **parts** do not. In any case, denying that the **system** understands Chinese is not the same as denying that the **human** in the room understands Chinese.

This 'systems reply' is often supplemented by a 'robot reply', a version of the Total Turing Test in which the participant memorizes the instruction book and is endowed with other kinds of cognition, such as those involving sensors (e.g., perception) and effectors (e.g., acting). Finally, as in the AI-complete situation, perhaps a suitably rich sort of symbol manipulation (combining linguistically derived concepts with perceptually derived concepts) could provide the

apparently missing semantic component (Rapaport, 2000).

**Is Cognition Subjective?** Both Block's and Searle's objections also raise the issue of whether cognition (or natural-language understanding) is a subjective phenomenon: a participant passes a Turing Test **with respect to a given interrogator.** Thus, a given lookup table might not work for a given interrogator, and different interrogators might disagree on whether a participant in a Chinese Room demonstrates understanding.

## Is Cognitive Behavior Really Cognition?

A participant will either pass the Turing Test or not. If **no** computer ever passes it – that is, if only humans (or other entities independently judged to be capable of cognition, such as extraterrestrials) pass it – then the Test succeeds in the sense that it offers a (behavioral) mark of (human) cognition. If, on the other hand, a computer **does** pass it, that could be for two reasons: either it is **really** cognitive (and able to **convince** the interrogator of this), or else it is not but only **appears** to be.

One way of phrasing the fundamental problem posed by the Turing Test is whether mere cognitive behavior **is** real cognition. There are three possibilities: (1) Real cognition might be 'mere' input-output behavior (statistical correlation). (2) Or real cognition might be input-output behavior **plus** any intervening process (including the 'Aunt Bertha' kinds of table-lookup processes or Eliza-like natural-language-processing tricks). (3) Or real cognition might be input-output behavior with an intervening process **of a 'cognitive' kind**, such as a reasoning process, a symbolic computational process, a connectionist computational process, or a 'dynamic systems' process, implementing some psychological theory.

The Turing Test does not seem to distinguish among these three possibilities. Should it, i.e., is it only external behavior that counts as cognition, or is the internal process (if any) important, too (as Shannon and McCarthy, 1956 suggested)? Consider a 'magical' natural-language-processing machine: a pure input-output machine that takes natural language input (from the interrogator) and 'magically' yields (appropriate) natural-language output (i.e., a 'function machine' with no internal processes). Suppose this magical machine passes a Turing Test. Is it cognitive? Turing might have argued that it was, even though we could not know why or how it managed to understand natural language (after all, it's magic!). On the other hand, although we know that our brain is a non-magical natural-language processor, we are not yet at the stage where we can say with certainty that we know why or how **it** manages to understand natural language.

Thus, the fundamental problem of the Turing Test is the classic problem of intension vs. extension: Is the **way** in which a result is obtained or an act is performed (i.e., the result or act viewed intensionally) more important than the **result** itself (i.e., the result or act viewed extensionally)? From the perspective of linguistics, is the ability to respond to stimuli with 'fluent' verbal behavior that is indistinguishable from 'normal' human verbal behavior different from a 'deeper', 'cognitive' ability of 'understanding language'? (This is nearly the debate between B. F. Skinner's theory of verbal behavior and Noam Chomsky's theories of grammar.)

Thus, there are two possibilities as to why a participant might pass a Turing Test: its cognitive behavior is (merely) extensionally equivalent to (real, human) cognition, or its cognitive behavior is (also) intensionally equivalent to real (human) cognition. Should the first possibility vitiate the Test? Suppose that the interrogator (an external observer) takes the 'intentional stance' and ascribes cognition to the participant. It does not follow that the participant is **not** really (intensionally) cognitive. After all, one reason that ascribing an internal, intensional cognitive state or process might be the best explanation for an entity's behavior is that the entity really produced its behavior via that process.

Consider the Aunt Bertha memorization machine. Suppose that the table lookups can be done in real time. We seem to be faced with two choices: assert that it is a counterexample to the Turing Test or accept that the Aunt Bertha machine really is cognitive (and not merely **behaving** cognitively). On the side of real cognition, we could argue that the Aunt Bertha machine would not work – that only a real cognitive agent could pass the Turing Test. The Chinese-Room Argument would **not** be a counterexample in this case, because the Chinese-Room system consisting of the human together with the instruction book might only pass if the instruction book were AI-complete, containing a full computational cognitive theory. To assert that it could not contain such a theory is to go beyond the empirical evidence, especially because we do not have that evidence yet.

Or, also on the side of real cognition, we could bite the bullet and say that the Aunt Bertha machine **is** cognitive **despite** being a mere lookup table, i.e., that appropriate lookup tables are just one way of being cognitive (cf. Putnam's (1960) Turing-machine

lookup-table analogy to the human mind). In other words, the **appearance** of cognition **is** real cognition; i.e., the external ascription of cognition is all that matters (as in the pen pal case). Arguably, our ordinary notion of 'understanding' applies to an omniscient deity that would never have to draw inferences or in any way 'process' information. Thus, if table lookup is not cognition, then such omniscience would not be cognitive, which seems to violate our pre-theoretical intuitions.

Consider an analogy: a Turing Test for seeing. Under what conditions would it be fair to say that an external observer can determine whether an entity sees? Is mere 'seeing behavior' enough?, e.g., is it enough that the entity reacts appropriately to visual stimuli whether or not it has any internal representation of them? Plants not only respond tropistically to light, but there is even some evidence that they can sense the presence of distal stimuli by means of light reflection. So, do plants see? Probably not, in the sense that they probably have no internal representation of the external stimuli – but can that be determined behaviorally? Consider 'blindsight': there is evidence that some blind humans can respond appropriately to visual stimuli. Do they see (despite their protestations that they cannot)? This case is especially puzzling, because it is conceivable that such blindsighted humans have internal representations that they use in processing the external stimuli, yet they are not aware of them.

## Conclusion

Turing hoped that his Test would replace the question, 'Can machines think?' It has not – in large part because of disagreement over whether it is an appropriate replacement. Any formal idea that is proposed to replace an informal one is bound to suffer from this problem, for the formal idea will always include some feature that the informal one appears to exclude, and vice versa. In the case of the Turing Test, this has given rise to the debate over its sufficiency for thinking and for cognition in general. Nevertheless, more than 50 years after its appearance, the Turing Test continues to help focus debate on issues concerning the relation of cognition to computation.

*See also:* Chomsky, Noam (b. 1928); Descartes, René (1596–1650); Natural Language Processing: Overview; Natural Language Processing: System Evaluation; Natural Language Understanding, Automatic; Putnam, Hilary (b. 1926); Text Retrieval Conference and Message Understanding Conference; Turing, Alan Mathison (1912–1954).

## Bibliography

Ahn L von, Blum M & Langford J (2004). 'Telling humans and computers apart automatically.' *Communications of the ACM 47(2) (February)*, 56–60.

Akman V & Blackburn P (eds.) (2000). 'Special issue on Alan Turing and artificial intelligence.' *Journal of Logic, Language and Information. 9(4)*.

Argamon S, Koppel M, Fine J *et al.* (2003). 'Gender, genre, and writing style in formal written texts.' *Text 23(3)*, 321–346.

Batacharia B, Levy D, Catizone R *et al.* (1999). 'CONVERSE: a conversational companion.' In Wilks Y (ed.) *Machine conversations*. Boston: Kluwer Academic Publishers. 205–215.

Block N (1981). 'Psychologism and Behaviorism.' *Philosophical Review 90(1)*, 5–43.

Bradford P G & Wollowski M (1994). 'A formalization of the Turing test.' *Technical Report 399*. Bloomington, IN: Indiana University Department of Computer Science.

Colby K M (1981). 'Modeling a paranoid mind.' *Behavioral and Brain Sciences 4*, 515–560.

Colby K M, Hilf F D, Weber S *et al.* (1972). 'Turing-like indistinguishability tests for the validation of a computer simulation of paranoid processes.' *Artificial Intelligence 3*, 199–221.

Dennett D C (1968). 'Intentional systems.' *Journal of Philosophy 68*, 87–106.

Descartes R (1637). *Discourse on the method of rightly conducting the reason and seeking for truth in the sciences*. In Haldane E S & Ross G R T (trans.) (1970). *The philosophical works of Descartes*, vol. I. Cambridge, UK: Cambridge University Press. 79–130.

Harnad S (1989). 'Minds, machines and Searle.' *Journal of Experimental and Theoretical Artificial Intelligence 1(1)*, 5–25.

Hirsch E D Jr (1987). *Cultural literacy: what every American needs to know*. Boston: Houghton Mifflin.

Lenat D B (1995). 'CYC: a large-scale investment in knowledge infrastructure.' *Communications of the ACM 38(11)*, 33–38.

Loebner H G (1994). 'In response [to Shieber (1994a)].' *Communications of the ACM 37(6)*, 79–82.

Loebner H G (2003). 'The first Turing test.' *Loebner Prize*. http://www.loebner.net/Prizef/loebner-prize.html.

Moor J H (ed.) (2000–2001). Special issues on the 'Turing test: past, present and future.' *Minds and Machines 10(4)*, 11(1).

Moor J H (ed.) (2003). *The Turing test: the elusive standard of artificial intelligence*. Dordrecht, The Netherlands: Kluwer Academic Publishers.

Putnam H (1960). 'Minds and machines,' In Hook S (ed.) *Dimensions of mind: a symposium*. New York: New York University Press. 148–179.

Rapaport W J (1988). 'Syntactic semantics: foundations of computational natural-language understanding.' In Fetzer J H (ed.) *Aspects of artificial intelligence*. Dordrecht, Holland: Kluwer Academic Publishers. 81–131.

Rapaport W J (2000). 'How to pass a Turing test: syntactic semantics, natural-language understanding, and

first-person cognition.' In Akman & Blackburn (eds.). 467–490.
Rey G (2002). 'Searle's misunderstandings of functionalism and strong AI.' In Preston J & Bishop M (eds.) *Views into the Chinese room: new essays on Searle and artificial intelligence*. Oxford: Oxford University Press. 201–225.
Sato Y & Ikegami T (2004). 'Undecidability in the imitation game.' *Minds and Machines 14(2)*, 133–143.
Searle J R (1980). 'Minds, brains, and programs.' *Behavioral and Brain Sciences 3*, 417–457.
Shannon C E & McCarthy J (1956). 'Preface.' In Shannon C & McCarthy J (eds.) *Automata studies*. Princeton, NJ: Princeton University Press. v–viii.
Shieber S M (1994a). 'Lessons from a restricted Turing test.' *Communications of the ACM 37(6)*, 70–78.
Shieber S M (1994b). 'On Loebner's lessons.' *Communications of the ACM 37(6)*, 83–84.
Shieber S M (ed.) (2004). *The Turing test: verbal behavior as the hallmark of intelligence*. Cambridge, MA: MIT Press.
Shieber S M (2005). 'The Turing test as interactive proof.' http://www.eecs.harvard.edu/shieber/Biblio/Papers/turing-interactive-proof.pdf.
Turing A M (1936). 'On computable numbers, with an application to the Entscheidungsproblem.' *Proceedings of the London Mathematical Society*, Ser. 2, *42*, 230–265.
Turing A M (1950). 'Computing machinery and intelligence.' *Mind 59*, 433–460.
Weizenbaum J (1966). 'ELIZA – a computer program for the study of natural language communication between man and machine.' *Communications of the ACM 9*, 36–45.
Weizenbaum J (1967). 'Contextual understanding by computers.' *Communications of the ACM 10*, 474–480.

# Turing, Alan Mathison (1912–1954)

**B Sypniewski**, Rowan University, Camden, NJ, USA



Alan Turing (1912–1954), one of the most important logicians and mathematicians of the 20th century, developed the basic theory for computer science and participated in the breaking of the German secret codes during World War II. Turing's contributions to linguistics usually go unnoticed. As an abstract thinker, his achievements crossed disciplinary boundaries.

In 1937, Turing published a paper usually referred to by its shortened title, 'Computable numbers,' in which he described what is now known as a Turing machine. A Turing machine is an abstract machine that executes a program (a set of instructions) one step at a time. In theory, a universal Turing machine can compute anything that is computable. The Turing machine is the abstract basis for modern digital computers.

In 1950, he published another important paper, 'Computing machinery and intelligence,' which raised the question of whether a machine, specifically a digital computer, could think. The paper described what has come to be known as the 'Turing test,' in which an interrogator questions a hidden human and computer. All questions and answers are communicated by typewritten message. The interrogator's job is to distinguish between the computer and the human solely on the basis of the answers. One player tries to deceive the interrogator; the other player tries to help the interrogator. Turing's argument was that if the computer successfully played what he called the 'Imitation Game,' it could be said to think, since it effectively mimicked a human being.

A grammar may be considered a Turing machine. Generative grammar is claimed to be a method of generating 'correct' and only 'correct' sentences in a language. The Turing test is a test of intelligence that equates intelligence with language use. While Turing held out high hopes for successfully programming a computer to pass his test (Turing test contests have been held, and at least one computer is claimed to have passed), no one has yet shown that language use is computable in the mathematical sense. Turing showed that if intelligence equates with some level of language use, language use cannot be limited to grammar. Taking his two great accomplishments together, we can say that even if grammar is shown to be a universal Turing machine, simply creating 'correct' sentences in a language is insufficient to pass the Turing test. Turing's achievements question whether the focus of linguistics should be restricted to grammar.

An interesting development in Turing's thoughts occurred between his two great papers. A Turing machine is very simple and has the ability only to react. In the last section of his 1950 paper, Turing no longer talked about a machine that is programmed but about one that learns, perhaps by a mechanism akin to natural selection. In between his papers, Turing helped to create Britain's first computer and had daily experience with it. Turing's theoretical computers responded to conversations, not equations.