# The Turing Test and Alternate Rationalities

Mete Atamel
Department of Philosophy, Gettysburg College, Gettysburg, PA 17325, USA;
E-mail: meteatamel@yahoo.com

**Abstract.** Since the introduction of the imitation game by Turing in 1950, there has been much debate about the possibility of machine intelligence and the accountability of the imitation game as a test of intelligence. Despite its limitations as a test of intelligence, the imitation game is a small yet important step in freeing Artificial Intelligence from the unnecessary constraint of imitating human intelligence. Viewing human intelligence as one possible form of intelligence will enable AI researchers to seek alternate rationalities that will create richer ways of understanding intelligence.

**Key words:** Artificial Intelligence, Imitation Game, Turing Test, Chinese Room Experiment, Searle, Total Turing Test.

## 0. Introduction

Can machines think? The British mathematician Alan Turing tried to answer this question more than fifty years ago. His "Computing Machinery and Intelligence" paper (Turing, 1950) not only gave birth to a new field called Artificial Intelligence (AI), but it also fueled a great deal of debate among philosophers on the nature of the mind and intelligence.

Turing proposed his famous imitation game as a replacement for the question “Can machines think?” His intention was to bypass the mystery of intelligence by equating perplexing internal brain functions with more comprehensible outward behaviors. Turing’s functionalist approach to understanding the mystery of intelligence may initially be considered too simplistic and even useless. However, careful consideration of the imitation game reveals that in his functionalist account, Turing set the right direction for Artificial Intelligence, giving greater freedom for the development of modes of artificial rationality distinct from human rationality.

Instead of judging the validity of Turing’s imitation game as a test of intelligence, it is more useful to consider it as an attempt to liberate intelligence from the monopoly of human intelligence. Following Turing’s direction will provide us ample opportunity to understand the nature of intelligence, whether in human or alternate forms.

## 1. The Imitation Game and Its Implications

As the father of Computer Science and a talented mathematician, Turing passionately believed that intelligence was within the scope of computable operations. The first indication of his belief in machine intelligence goes back to as early as the end of World War II. In one of his statements in 1945 he says:

> Given a position in chess the machine could be made to list all the 'winning combinations' to a depth of about three moves in either side. This is unlike the previous problem, but raises the question 'Can the machine play chess?' It could be fairly easily made to play a rather bad game. It would be bad because chess requires intelligence. We stated at the beginning of this section [i.e. when describing how programming is done] that the machine should be treated as entirely without intelligence. There are indications however that it is possible to make the machine display intelligence at the risk of its making occasional serious mistakes. By following up this aspect the machine could probably be made to play very good chess. (Hodges, 1999, pp.30)

This statement indicates the path that Turing would follow five years later. In 1950 (Turing, 1950), Turing proposed a test called the imitation game which is also known as the Turing Test (TT). The game is initially played by three humans --a man, a woman and a judge. The only communication among the judge, the man and the woman is via teletype machine. The judge has to determine the gender of the person at the other end by only asking questions via teletype. But both of the players are trying to convince the judge that they are female. Now, imagine that one of the players is replaced by a machine and the task of the judge is changed. He now must distinguish between the human player and the machine player rather than distinguishing between genders. If the judge cannot make the distinction between the human and the machine player, then the machine is said to have passed the test. In other words, if the machine fools the judge to believe that it is a human, the machine passes the test.

According to Turing, passing the imitation game is enough to conclude that the machine is sufficiently like the man to be accorded the same intellectual status. In other words, "Can a machine think?" can be reduced to "Can a machine pass the imitation test?" The absence of a general definition of intelligence at the time enabled Turing to create a definition that further helped him to defend his intelligence test. He proposed "looks like a duck, walks like a duck, quacks like a duck, it's a duck" as the definition of intelligence (Siebel). He believed that we shouldn't worry about what's going on inside of the machine. What matters is the outward behavior and if it is indistinguishable from the outward intellectual behavior of a human then it is, by his definition, intelligent (Siebel).

The validity of the imitation game as an ultimate test of intelligence is still disputable. Nevertheless, it is the first major attempt to treat machine intelligence (and by implication human intelligence) functionally, independent of considerations of the body. Turing's imitation game can be thought of as an attempt to revolutionize the classic notion of intelligence, liberating our concept of intelligence from the monopoly of human intelligence.

**2. Searle's Chinese Experiment & Simon on Computer Intelligence**

Turing’s bias towards functionalism and the deep reductionism of the imitation game raised opposition among philosophers of mind. Among those opposed to the TT, John Searle was probably the most influential one. In "Minds, Brains, and Programs," (Searle, 1980) Searle introduces his famous "Chinese room" experiment. Searle wants us to imagine a man sitting inside a closed room. As input, he gets a piece of paper covered with Chinese symbols (squiggly marks as he refers to them) through a hole in the door. The man then looks at his big reference book that contains all the rules to match certain squiggly marks with other squiggly marks. So according to what input he gets, he looks at his book, finds the corresponding marks, writes them down and presents those marks as output from the hole of the door. Searle's point is that the man in the room can pass the TT. He can play the imitation game with a native Chinese speaker and he can fool the native Chinese speaker into thinking that he can speak Chinese. In truth, he has no understanding of Chinese whatsoever. All he does is matching inputs with corresponding outputs without any understanding. Searle takes this idea and concludes that intelligence is not simply a matter of computation or symbol manipulation as Turing suggests.

Searle not only refuses Turing's computational view of mind but he also claims that the imitation game is not really a test of intelligence. As convincing as it sounds, Searle's Chinese room experiment hasn't gone unchallenged. The most convincing reply, also known as the systems reply, came from computer scientists at Berkeley. According to the systems reply, although there is no understanding at the level of man in the room, there is understanding in the whole system. They claim that the room combined with the man and the rule book possesses some understanding. We don't deny understanding from a brain because of one single neuron's lack of understanding. So why should we deny understanding from the whole system because the man does not understand Chinese? Searle responds to the systems reply by letting the man memorize the rule-book and do all the manipulations in his head. Now the man is the system and he still doesn't understand. A possible answer to Searle's reply is following. “If it were really possible for the man to memorize all the rules and do all the manipulations in his head, then there could be a ‘passenger’ personality, an entity that is using the man's brain but is not the man, and who exists only because the man is doing the calculations prescribed by the rule-book” (Siebel).

Although it runs contrary to all intuition, Turing's imitation game is not so easy to refute. If a machine talks like a human, acts like a human, then on what grounds can it be denied of intelligence? Turing's imitation game is so basic that perhaps we feel reluctant to identify intelligence so simply. If intelligence can be identified functionally, how can we claim that we have a complex brain that produces miraculous and mysterious results such as reasoning and creativity? As Herbert Simon points out in his article "Machine as

Mind" “perhaps we are reluctant to give up our human uniqueness--of being the only species that can think big thoughts. Perhaps we have ‘known’ so long that machines can't think that only overwhelming evidence can change our belief.” (Simon, 1995, pp. 23)

In his article, Simon portrays minds as machines. He explains how seemingly mysterious aspects of our minds, such as selective heuristic search, recognition, semantics and insight problems, are actually not as mysterious as they seem. He draws on significant data from psychology to support his ideas. He claims that words like "intuition" or "creative" are not as human as they seem to be. For example, he defines "intuition" as problem solving by recognition and it can be easily modeled by production systems. He also talks about a "creative" computer program called BACON, which, when given the data available to scientists in historically important situations, has rediscovered Kepler's Third Law, Ohm's Law, Boyle's Law and many others. The following paragraph from his article is very interesting.

> It has been argued that a computer simulation of thinking is no more thinking than a simulation of digestion is digestion. The analogy is false. A computer simulation of digestion is not capable of taking starch as an input and producing fructose or glucose as outputs. It deals only with symbolic quantities representing these substances. In contrast, a computer simulation of thinking thinks. It takes problems as inputs and (sometimes) produces solutions as its outputs. It represents these problems and solutions as symbolic structures, as human mind does (Simon, 1995, pp.24)

Turing and Simon might be right. Maybe the human mind does not reach its goals mysteriously or miraculously. As Simon suggests, maybe “even its sudden insights and "ahas" are explainable in terms of recognition processes, well-informed search, knowledge-prepared experiences of surprise, and changes in representation motivated by shifts in attention.” Maybe the unexplainable is explained once we accept the fact that computers have been thinking "creatively" and "intuitively" for the past 35 years.

### 3. DeLancey's Passionate Engines

Even if we accept the "creative" and "intuitive" thinking of computers, one question still remains unanswered. Is the TT really a test of intelligence, and if so, of human intelligence or intelligence more generally? Before answering this question, we need to define what we mean by intelligence and intelligent behavior.

In his book *Passionate Engines* (Delancey, 2002), Craig DeLancey investigates the relationship between what emotions reveal about the mind and artificial intelligence. He shows that our best understanding of the basic emotions provides essential insights into key issues in the philosophy of mind and AI. He believes that emotions and behaviors are not completely unrelated. In fact, he asserts that emotions are inseparable from actions, which is why he refers to emotions as effects in his book. He offers new ways to understand the mind, suggesting that autonomy--and not cognition—should be the core problem of the philosophy of mind and AI.

When he reflects on the computational theory of mind, DeLancey claims that symbolic computational functionalism cannot account for the effects of some basic emotions. Hence he believes that functionalism as a theory of human and other animal minds is inadequate. In "Affective Engineering" a key section of his book, he reasserts the importance of autonomy and how it should be the central focus of AI. Delancey asserts that the idea of starting with a pure symbol or proposition manipulation, as modeled by a symbol manipulating language, and getting autonomous behavior out of this has been a failure.

He suggests viewing an intelligent system not as a mere symbol manipulator but foremost an autonomous passionate engine. We have little reason to think of the cognitive abilities that humans have as anything but one among many strategies for autonomy. Cognition is neither necessary nor sufficient for autonomy. His central thesis is the body and the mind, passion and action, are inseparable.

DeLancey's argument is very convincing because his analysis of mind situates mind in the context of nature and evolutionary history. Instead of trying to reduce intelligence to some sort of function or automata, he argues that we ought to enrich our understanding of intelligence by viewing it as manifested through the body. His introduces autonomy as a necessary condition for intelligence and supports his thesis by pointing out that many nonhuman animals reveal a high degree of autonomy to such an extent that they differ from us only in lacking comparable cognitive skills. He also believes that we have little reason to think of the cognitive abilities that humans have as anything but one among many strategies for autonomy. He links brain and body together so tightly that brain without body and intelligence without autonomy is unthinkable.

Delancey provides a new path for AI and following this path can *only* increase our knowledge and understanding of the potentiality of AI. Nevertheless, Delancey’s conception of embodied intelligence might also be thought to restrict AI. As Delancey argues in his book, he expects to see robotics growing cheaper and more powerful in the future, and hence an essential part, if not the core, of AI in the future (Delancey, 2002). Delancey probably does not consider Deep Junior, a machine that won three chess games against the best human chess player, intelligent since it does not have a body. This approach might lead us to think that having a body is a necessary condition for intelligence. If AI’s aim is to imitate human intelligence, then requiring a body for intelligence might be a necessity. However, if we want to simulate or create forms of intelligence in general, requiring a body for every intelligent agent will not help AI.

Passionate Engines completely eliminates the possibility that the TT can be an intelligence test for machines because the TT not only rests on the assumption that cognition is sufficient for intelligence but it also disregards the importance of autonomy and feeling in intelligence.

**4. Is Turing Test Really a Test of Intelligence?**

DeLancey brought autonomy and the inseparability of mind and body into efforts to define intelligence, and pointed to biology as a proper guide for AI rather than physics or mathematics. Embracing DeLancey's argument leaves no open doors for the TT. Given the truth of Delancey's argument and its compatibility with biology and evolution, it becomes obvious that the TT is not sufficient as a test of intelligence because it completely ignores the importance of autonomy and it marks a major separation between the mind and the body.

Since Turing’s test completely ignores the importance of autonomy, does it mean that it is completely useless? The answer is not straightforward. As a complete test of intelligence, the Turing test can be considered useless. However, its power as an indication of intelligence cannot be ignored.  The power of Turing's test comes from the fact that it is based on language and language is a powerful and subtle tool of human intelligence. Since language only emerges from some sort of intelligence, we can be tempted to regard the TT as an intelligence test because what the TT does is basically detecting language.

But we shouldn't so easily let the trap of language fool us into thinking that the TT is a test of intelligence. Biological evolution seems to suggest that robotic capabilities come before linguistic ones. There are many species with autonomous, robotic capabilities with no linguistic capacity, but none that has linguistic capacity without robotic capacity. We don't deny the limited intelligence of a chimpanzee because of its lack of linguistic capabilities although it is obvious that a chimpanzee can't come close to passing the TT. It is hard to imagine that a TT candidate could chat with you coherently about the objects in the world without even having encountered any objects directly (Harnad, 1991).

The TT focuses too narrowly on cognition, disregarding other bodily functions that are essential for intelligence. Presenting the cognitive abilities of intelligence as the representation of an entire organism's intelligence fails to capture intelligence. Autonomy combined with feelings makes possible an organism's direct experience of world. Without direct experience and related feelings, language is nothing more than a string of characters. The TT has no implications for autonomy or feelings so it is hard to be regarded as a test of language let alone a test of intelligence.

Despite the limitations of the TT in measuring intelligence, Turing had the right idea when he first proposed the imitation game. If we want to measure intelligence, it has to depend on something that we can actually measure. Turing proposed language as a measure of intelligence because language is something that we can identify. Although it turned out that language is not powerful enough to identify intelligence, this does not mean that the imitation game is wrong in essence. There are many other aspects of human intelligence. The imitation game can be taken as a functional basis for appraising intelligence and upgraded to provide a more complete test of intelligence using other aspects of human intelligence.

**5. A Total Turing Test**

When Turing first proposed the imitated game, his main idea was "if it looks like a duck, walks like a duck, quacks like a duck, it's a duck". He was right that if an "artificial" duck acts like a duct in every respect, we have no reason to doubt its authenticity. However the test that he suggested for human intelligence did not actually test if the machine acts like a human in every respect. The imitation game might be assumed to be testing only "quacks like a duct" although I'm not even sure if typing "quack" on a teletype machine with no feelings involved is the same thing as an actual "quack" sound coming from an actual duck as it is of fear or joy. The TT provides no test for "looks like a duck" or "walks like a duct". With AI, we are not so much concerned about the "looks" of an artificial being. So "looks like a duct" is not that important for us. However, we should be concerned about "walks like a duct". In other words, we should be concerned about intelligence as autonomy. An intelligence test has to account for autonomy if it is seriously be considered as a test.

In "Other Bodies, Other Minds," Stevan Harnad proposes a more stringent test called the Total Turing Test (TTT) in place of Turing's original "pen-pal" version of the TT. To pass Harnad's Total Turing Test, "The candidate must be able to do, in the real world of objects and people, *everything* that real people can do, in a way that is indistinguishable to a person from the way real people do it" (Harnad, 1991)

Harnad's TTT is much stronger than the TT. First, it requires that a candidate is an actual physical object such as a robot. Second, passing the TTT entails passing the TT. Besides the TTT requires full robotic capabilities and this makes the successful linguistic performance more probable because the machine now has a direct relationship with the world. Most importantly, the TTT is immune to Searle's Chinese room experiment because the only way to actually perform the Chinese room experiment would be to be the actual machine in question. Moreover, the sensorimotor grounding of the TTT further eliminates "the mindless symbol manipulator" argument.

It seems like Turing's "looks like a duck, walks like a duct, quacks like a duck, it's a duck" argument is fully achieved by the TTT. The only objection to the TTT can be directed from a "Quality of Feelings" angle. As Frank Jackson points out in his article "Epiphenomenal Qualia," no matter how much we learn about the brain, its related states, their functional role, and so on, we don't know anything about the hurtfulness of pains, the itchiness of itches, the smell of a rose and so on unless we experience those feelings ourselves (Jackson, 1984). This means that physicalism leaves out something. As Harnad himself acknowledges, the TTT is incapable (as TT) of distinguishing systems that really have "*private experiences*" that we each know *exactly* what it's like to have from systems that might "*just be behaving exactly as if they had a mind without experiencing anything*".

This argument is in fact no different from assuming that no one else has a mind. There is no evidence for Person A that Person B has a mind unless Person A can experience what Person B experiences at a certain time and conclude that Person B's experiences are identical to his experiences. Since it is impossible to experience what other people can experience, it is impossible to be completely sure that another person has a mind. In this regard, it is unfair to deny experience from a machine because we assume that it cannot experience what we experience. Even if we are sure that it cannot experience what we experience, we should still leave an open door for some type of a machine experience because after all the only way to know how it is to be a machine is to be a machine.

**6. How Useful is the TTT?**

Isn't the TTT essentially building a human? If we set ourselves the goal of creating some entity that acts like a human in every respect, aren't we essentially saying that our goal is to create a human being out of metal pieces? Isn't the philosophical value of AI to understand the source or nature of intelligence through simple, somewhat intelligent and rational machines? It seems that by requiring a TTT type of test as the absolute intelligence test for our agents, we are getting beyond the initial premise and purpose of AI.

One question to consider is "Why do the Turing Test and the Total Turing Test try to test the human capabilities of a machine?" For example, the Turing Test tries to test the linguistic capabilities of a machine. Linguistics is only one aspect of human behavior and it is exclusive to humans. On the other hand, the TTT tries to test *every aspect* of human behavior. I believe that both tests derive from our constant urge to create a version of ourselves through AI. Since these types of tests determine the direction of AI research, we are always bound to imitating human intelligence.

One of the biggest problems in AI is that we try to measure the success of AI by how much a machine can imitate human behavior. That's why we initially had the TT and that's why it turned out to be insufficient to account for intelligence. Then we turned to TTT that essentially requires building a complete human. TTT can be considered a complete test of intelligence but it does not provide new ways of creating and understanding intelligence.

If we are trying to understand the source of intelligence and rational behavior, we shouldn't be bound by the goal of imitating human intelligence and human behaviors only. We have to accept that intelligent and rational behaviors can be created in ways that do not use human intelligence as their model. AI’s success should be measured in terms of the production of rationality and intelligence in machines without feeling constrained to imitate human intelligence.

**7. Alternate Rationalities**

Imitating human intelligence can be beneficial as a field in AI but it shouldn’t be the central motivation of AI. Despite its limitations, the Turing Test was one small step to isolate intelligence from consideration of the body. Now we need to free AI from the unnecessary constraint of imitating human intelligence. Turing’s emphasis on outward behavior as a measure of intelligence has to be taken seriously no matter how limited his conception of intelligence was. We have to accept that intelligence can be obtained in endless different forms and that internal operations performed to obtain outward intelligible actions can be different. Every rational and intelligent behavior exercised by a machine should be treated as another step in understanding mind whether in human form or in alternate machine form.

For a long time we wondered about flying. We always wanted to fly and we endlessly tried to imitate flying patterns of birds. Imitating birds helped us understand flying in general. But we could only fly after we discovered the fundamental physical concepts behind flying and created planes that obeyed those rules without worrying too much about one instance of flying, that is birds’ flying.  The same analogy applies to AI as well. We do not have to imitate or understand every aspect of human intelligence in order to understand or create intelligence. It’s time to accept the fact that Deep Junior, the computer that challenged Kasparov, the best human chess player, might possess a type of intelligence. Maybe Deep Blue does not possess the creative, intuitive and humanly type of intelligence that we perform while playing chess but it performs high speed search techniques that no human can perform and finally produces intelligible and rational chess moves quite often better than a human does. Instead of labeling Deep Blue as a mere calculator, it is much more constructive to analyze and try to understand the type of intelligence that it exhibits.

## Conclusion

As long as we try to understand intelligence by trying to imitate human intelligence, we will always be bound by the unrealistic expectation of creating human intelligence out of metal pieces. We can either try to imitate every single neuron in our brain and hope these neurons will end up working the same way that our brain does, use the TTT to test if our machines are “human” enough, or we can accept that intelligence can be obtained in alternate forms and try to acknowledge and create those alternate forms in order to better understand intelligence.

In his famous imitation game, Turing separated the bodily features of a human from human intelligence. In doing so, he initiated the idea that intelligence can exist independent of the human body. We should hurry up and follow in the direction that he pointed fifty years ago. It is time to accept that human intelligence is only one form of intelligence and that alternate rationalities can be created in alternate forms from the most familiar human form.

**References**

Colburn, T. (2000), *Philosophy and Computer Science*, New York, NY: M.E. Sharpe.

Delancey, C. (2002), *Passionate Engines*, New York, NY: Oxford University Press, Inc.

Harnad, S. (1991), 'Other bodies, Other Minds: A Machine Incarnation of an Old Philosophical Problem', *Minds and Machines* 1, pp. 43-54.

Hauser, L. (1993), 'Reaping the Whirlwind: Reply to Harnad's "Other Bodies, Other Minds"', *Minds and Machines* 3, pp. 199-226.

Hodges, A. (1999), *Turing: A Natural Philosopher*, New York, NY: Routledge.

Jackson F. (1984), 'Epiphenomenal Qualia', *Philosophical Quarterly*, 32, pp.127-136.

Penrose, R. (1989), *The Emperor's New Mind*, Oxford UK: Oxford University Press.

Rose, S. (1998), *From Brains to Consciousness*, Princeton NJ: Princeton University Press.

Searle, J.R. (1980), 'Minds, brains and programs', *Behavioral and Brain Sciences* 3, pp. 417- 457.

Seibel, P., "Turing Test, Etc.", http://www.ryerson.ca/~dgrimsha/courses/cps721/turing.htm

Simon, A. (1995), 'Machine As Mind' in *Android Epistemology*, Ford, K., C. Glymour and P.Hayes, eds, American Association for Artificial Intelligence, MIT, Cambridge, MA: The MIT Press.

Turing, A.M. (1950), 'Computing Machinery and Intelligence', *Mind* 59, pp.433-460.