

# Taking the fifth amendment in Turing's imitation game

Warwick, K. and Shah, H.

**Author post-print (accepted) deposited by Coventry University's Repository**

**Original citation & hyperlink:**

Warwick, K. and Shah, H. (2016) Taking the fifth amendment in Turing's imitation game. *Journal of Experimental & Theoretical Artificial Intelligence*, volume 29 : 287-297  
<http://dx.doi.org/10.1080/0952813X.2015.1132273>

DOI 10.1080/0952813X.2015.1132273

ISSN 0952-813X

Publisher: Taylor and Francis

**This is an Accepted Manuscript of an article published by Taylor & Francis in *Journal of Experimental & Theoretical Artificial Intelligence* on 7<sup>th</sup> January 2016, available online: <http://www.tandfonline.com/10.1080/0952813X.2015.1132273>**

**Copyright © and Moral Rights are retained by the author(s) and/ or other copyright owners. A copy can be downloaded for personal non-commercial research or study, without prior permission or charge. This item cannot be reproduced or quoted extensively from without first obtaining permission in writing from the copyright holder(s). The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the copyright holders.**

**This document is the author's post-print version, incorporating any revisions agreed during the peer-review process. Some differences between the published version and this version may remain and you are advised to consult the published version if you wish to cite from it.**

# Taking the Fifth Amendment in Turing's Imitation Game.

Kevin Warwick and Huma Shah

Coventry University, Priory Street, Coventry, CV1 5FB, UK

Email: [k.warwick@coventry.ac.uk](mailto:k.warwick@coventry.ac.uk) , [h.shah@coventry.ac.uk](mailto:h.shah@coventry.ac.uk)

Corresponding author: Kevin Warwick, tele: 44-247765-9893

**Abstract:** In this paper we look at a specific issue with practical Turing tests, namely the right of the machine to remain silent during interrogation. In particular we consider the possibility of a machine passing the Turing test simply by not saying anything. We include a number of transcripts from practical Turing tests in which silence has actually occurred on the part of a hidden entity. Each of the transcripts considered here resulted in a judge being unable to make the 'right identification', i.e. they could not say for certain which hidden entity was the machine.

**Keywords:** Deception Detection, Natural Language, Turing's Imitation Game, Chatbots, Machine Misidentification

## ACKNOWLEDGEMENTS

Figure 1: Harjit Mehroke for Figure 1a; C.D. Chapman for Figure 1b.

## INTRODUCTION

In his 1950 paper entitled 'Computing Machinery and Intelligence' (Turing, 1950), Alan Turing wished to consider the question, "Can machines think?" Rather than get bogged down by definitions of both of the words "machine" and "think" he replaced the question with one based on a much more practical scenario, namely his imitation game.

Turing himself described the game in these terms: "The idea of the test is that a machine has to try and pretend to be a man, by answering questions put to it, and it will only pass if the pretence is reasonably convincing. A considerable portion of a

jury, who should not be expert about machines, must be taken in by the pretence” (Copeland, 2004). In this case Turing spoke about a jury (possibly 12 people) rather than the ‘average interrogators’ he mentioned separately (Turing, 1950).

Central to Turing’s thesis is a machine’s ability to post appropriate responses to any questions. However, what we look at in this paper is what if the most appropriate response, (as deemed by an entity) to a particular question, is silence? What is the thinking nature of a machine when, rather than responding to an inappropriate or inane question, it does not answer it with an utterance? Why should a truly intelligent machine ingratiate itself with humanlike responses just to be considered human? Is not the truly Turing-intelligent machine the one that knows when and why to be silent?

So what we consider here is the role/effect of silence in Turing’s Imitation Game on the part of a hidden entity, especially when that entity is a machine. We then push this to the extreme by looking at the case when a machine is systematically silent in a sustained fashion and how this is reflected in the standard interpretation of functioning of the game. Our argument is supported by practical examples obtained from actual tests when machines have remained silent. These examples, which have (we suspect) generally arisen from a machine or communications failure rather than a decision on the part of a machine, have been included mainly to indicate how interrogators (the terms interrogators and judges are used here interchangeably to mean the same thing) respond in such circumstances.

## INTERPRETING TURING

There are many different interpretations of Turing’s imitation game and much controversy has arisen as to which of these, if any, was Turing’s own intended version (Moor, 2003). The vast majority appear to view the game in the form of what is commonly known as the ‘Standard Turing Test’ (Sterrett, 2000), and this is the interpretation taken here. It is a literal interpretation based essentially on what Turing actually said in his presentations and papers and without recourse to tangential connections and/or pure conjecture on what a paper’s author believes that Turing really meant to say. We acknowledge as examples of this, that some see it as being something to do with artistic and emotional intelligence (Smith, 2015), whereas others deem it to be concerned with modelling the human mind by generating its verbal performance capacity (Harnad, 2014).

Others meanwhile regard it in terms of considering the gender aspect, the sex of the human foil being important in the test (Sterrett, 2000; Lassègue, 1996; Hayes & Ford, 1995; Genova, 1994). In one interpretation of Turing’s test the female is expected to tell the truth, but we are not far off that time when silence was preferred to the “jabbering” of women, because “speech was the monopoly of man” and that “sounds made by birds were part of a conversation at least as intelligible and intelligent as the confusion of tongues arising at a fashionable lady’s reception” (Herbert Maxwell, 19<sup>th</sup> century Scottish politician, see Bourke, 2011).

In his 1950 paper Turing said: “I believe that in about fifty years' time it will be possible, to programme computers, ... , to make them play the imitation game so well that an average interrogator will not have more than 70 per cent chance of making the **right identification** after five minutes of questioning (Turing, 1950). To put this more simply, for a machine to ‘pass’ the Turing test, in all of the tests in which a machine takes part, the interrogators must make the wrong identification (i.e. not the right identification) 30% or more of the time after, in each case, 5 minute long conversations. However we can take it directly that the wrong identification is anything other than the right identification.

The Turing test involves a machine which pretends to be a human in terms of conversational abilities. The ‘right identification’ stated by Turing can either mean that a judge merely correctly identifies the machine or that they correctly identify, at the end of a paired conversation, which was the machine and which was the human (Traiger, 2000). These different viewpoints are discussed further later. However we are not so interested here with cases in which a judge mistakes a human for a machine. This phenomenon, known as the *confederate effect* (Shah and Henry, 2005), has been discussed elsewhere (Shah and Warwick, 2010a; Warwick et. al., 2013; Warwick and Shah, 2014). Such a decision might however affect the judge’s decision regarding the machine being investigated in parallel.

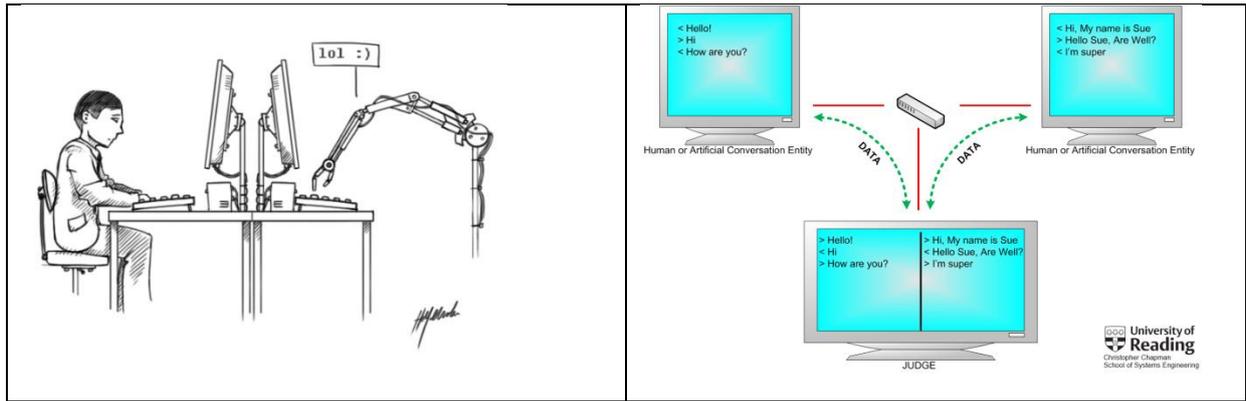
In its standard form, Turing’s imitation game is described as an experiment that can be practicalized in two different ways (see Figure 1) (Shah, 2011):

- 1) one-interrogator-one hidden interlocutor (Figure 1a),
- 2) one-interrogator-two hidden interlocutors (Figure 1b).

In both cases the machine must provide “satisfactory” and “sustained” answers to any questions put to it by the human interrogator (Turing, 1950: p.447). However, what about in the theoretical case when the machine *takes the 5<sup>th</sup> amendment*: “No person shall be held to answer”<sup>1</sup>! Would we grant “fair play to the machines”? (Turing, 1947; Hodges, 2010).

---

<sup>1</sup> Cornell University, Legal Information Institute: [http://www.law.cornell.edu/wex/fifth\\_amendment](http://www.law.cornell.edu/wex/fifth_amendment): “No person shall be held to answer for a capital, or otherwise infamous crime, unless on a presentment or indictment of a grand jury”



**Figure 1: Turing's two test for his imitation game: Left- Figure 1a, one-to-one; Right- Figure 1b, one judge-two hidden interlocutors**

Of the types of (standard) test looked at here, the 3-participant tests have previously been shown to be stricter tests, i.e. more difficult for machines, than 2-participant tests in which an interrogator converses with only one hidden entity, either a human or machine, at a time (Shah et. al., 2012). Hence these are the tests that we consider at length in this paper, particularly with the practical examples. It is worth making it clear however that for the main argument in the paper, the results apply to either type of practical test.

In the two hidden entity tests that the authors have conducted (see Figure 1b) judges were clearly told beforehand that in each parallel conversation one of the hidden entities was human and the other was a machine, the most recent example being in the 2014 Royal Society tests (Warwick and Shah, 2015). The interrogators/judges were however given no indication as to whether the LHS (Left Hand Side of the computer screen as they viewed it) or RHS would be human or machine. On the judges' score sheets each judge could mark both the LHS and RHS entities as being Human, Machine or they could say if they were Unsure about either or both (Shah et. al., 2012).

In the following section we consider a number of transcripts obtained from practical Turing tests. We refer here to 5 minute long tests only and show transcripts from such tests. Although this is the run time stated by Turing himself (Turing, 1950), it is not at all a critical issue with regard to the main argument raised in this paper. As you will see, in the tests carried out there was a hard cut off at the end of each discourse and no partial sentences were transmitted. Once a sentence had been transmitted it could not be altered or retracted in any way.

We have not edited the text of these transcripts in any way, indeed they are reported exactly as they occurred. Any typographical, spelling, punctuation or other grammatical errors that appear in the transcripts are those which actually appeared, they have not arisen through poor editorial practice on the part of the authors or publishers. Also any periods of silence or where no response appears are an exact indication of what actually happened.

## THE ROLE OF SILENCE IN TURING'S IMITATION GAME

In this paper we do not argue for or against a particular interpretation of the test, some of these have been explained earlier and in more depth elsewhere (Shah, 2013; Shah, 2011; Shah and Warwick, 2010a). What the authors contend in this paper is that Turing's ideas to examine machine thinking point towards numerous directions that are worthy of scientific enquiry, including the point we wish to raise here, which is to consider the question as to when would it be worthwhile for a machine to remain silent in an imitation game? Would that not be the strongest test for a machine, in an exploration of whether it can reply appropriately, to any and all inappropriate questions? Consider firstly this purely imaginary conversation (Transcript 1) taking place over 5 minutes between 10.58am and 11.03am one day between a human 'Judge' and a hidden interlocutor, 'Entity':

### Transcript 1

[10:58:14] Judge: Good day  
[10:58:30] Entity: A very good day to you too :)  
[10:59:01] Judge: How old are you?  
[10:59:22] Entity: That's a bit rude?  
[11:00:06] Judge: How old are you?  
[11:00:31] Entity: I'm not going to answer that, old/young it's irrelevant  
[11:00:47] Judge: How old are you?  
[11:01:27] Entity:  
[11:01:52] Judge: How old are you?  
[11:02:17] Entity:  
[11:02:45] Judge: How old are you?  
[11:02:52] Entity:

What if in that imaginary conversation the 'hidden entity' was a computer programmed with a personality of a middle-aged person who was sensitive about their age? What if that imaginary Turing test was extended to 11.08am and it continued for another five minutes with the hidden interlocutor refusing to answer, remaining silent and the Judge was not quite sure at the end of it whether they were talking to a human or a machine?

If the reader feels an interrogator would not in a realistic test simply repeat questions over and over then they would be incorrect. As an example, consider an actual transcript (Transcript 2) involving a human interrogator, Duncan Graham-Rowe, an award-winning science, technology, medical and environment journalist, acting as a judge in the 2006 Loebner Prize (Shah and Warwick, 2009).

### Transcript 2

Judge: Is it safe?  
Entity: Probably.

Judge: Is it safe?  
Judge: Is it safe?  
Entity: Do you think it's safe?  
Entity: Of course.  
Judge: Is it safe?  
Entity: Yes.  
Judge: Are you sure?

To note: An earlier utterance of “Is it safe?” appeared in the film ‘Marathon Man’ in which Laurence Olivier was quizzing Dustin Hoffman. Firstly we can confirm that the judge in Transcript 2 was human. Secondly the entity in this conversation was in fact a machine, Jabberwacky. Judges in practical Turing tests have introduced all sorts of their own tricks to ‘out the machine’, for example the 2005 Loebner Prize judge and philosopher Ned Block’s use of non-words “cslkjslw weicmslw weos” (Shah and Warwick, 2007)

Instances of no-reply at all from hidden interlocutors have actually occurred in practical Turing tests on a number of occasions, most recently in the series of experiments conducted at The Royal Society by both authors on 7 June 2014. Here is one of those conversations. The timing given is the exact timing on the day of the tests. The hidden entity in this case was the machine, ‘Cleverbot’. What follows is an actual five minute transcript (Transcript 3) in full:

#### Transcript 3

[10:58:08] Judge: good day  
[10:58:08] Entity:  
[10:58:46] Judge: is no response an answer  
[10:58:46] Entity:  
[10:59:35] Judge: am i not speaking you're language  
[10:59:35] Entity:  
[11:00:25] Judge: silence is golden  
[11:00:25] Entity:  
[11:01:32] Judge: shhh  
[11:01:32] Entity:  
[11:03:07] Judge: you make great conversation  
[11:03:07] Entity:

The hidden entity in this real Turing test conversation was not identified as being either human or machine, the Judge made the decision that they were ‘Unsure’. This was taken from a one judge-two hidden interlocutors, simultaneous comparison test (see Figure 1b). We have not shown the parallel hidden human entity transcript as it is not particularly relevant to the argument. Suffice to say that the hidden human, for which there was a corresponding fluid conversation, was categorised correctly by the judge to be a human. Despite this decision and initial instructions the interrogator remained firm that they were Unsure about the nature of the entity in Transcript 3.

Previously, in a series of Turing test experiments in 2008 at the University of Reading, on four occasions one machine recorded an ‘unsure’ classification from the Judge in its total of 12 tests – in fact that is an incorrect identification of greater than 30 per cent. We do not wish to say here, indeed it would be churlish of us to suggest, that that particular machine

'Brother Jerome' passed the Turing test, Turing's ideas are far too important to belittle them in this way. However following the Turing test argument rigorously, then that is exactly what happened.

The conversations under investigation between the same machine (Brother Jerome) and a different judge on each occasion took place in Session 2 on 12 October 2008 at Reading University (see Shah and Warwick, 2010abc). Again these were one judge-two interlocutor, simultaneous comparison tests as in Figure 1b. Three of the conversations are transcribed below, in each case the 'Judge' was the human interrogator and the 'Entity' was the machine, Brother Jerome. Different judges were involved in these transcripts:

#### Transcript 4

[12:34:03] Judge: wotcha, how's it going?  
[12:34:31] Judge: hello?  
[12:34:36] Entity:  
[12:34:40] Entity: Don't mind me - I'm just messing with your head :)  
[12:35:00] Judge: grr  
[12:35:31] Judge: how's your day been so far?  
[12:35:35] Judge: finding this interesting?  
[12:36:13] Judge: you still therE?  
[12:38:57] Entity:  
[12:39:22] Entity: So far but no further? What of the future, do you think?

In the case of Transcript 4, the Judge gave an 'Unsure' classification but also reported "test failed technical issues". This particular interrogator, an adult male and a consultant with Microsoft, was the most accurate judge in the 2008 series of Turing tests, in terms of the number of correct classifications of hidden entities. Importantly the judge did not classify the machine involved as being a machine, i.e. they did not identify the machine correctly.

#### Transcript 5

[13:12:39] Judge: Have you ever listened to mozrt? How does it make you feel?  
[13:13:11] Entity:  
[13:13:34] Judge: Did you write anything? I can't see it?  
[13:14:05] Entity:  
[13:14:50] Judge: Still not seeing anything, having computer trouble?  
[13:15:22] Entity:

The conversation, if it can be called that, concluded after 5 minutes. There were no further utterances from either party after the [13:14:50] judge input. The Judge in Transcript 5 was a male teenager. He made no other comment other than his 'Unsure' classification.

#### Transcript 6

[13:21:55] Judge: Hello.

[13:22:04] Entity:  
[13:22:29] Judge: Could I know the difference between a computer and a human from you?  
[13:23:02] Entity: Tell me more...  
[13:24:09] Judge: If you are a computer, you should tell me more.  
[13:24:41] Entity:  
[13:25:31] Judge: What you told me is a blank line.  
[13:26:03] Entity:  
[13:26:32] Judge: Hello??  
[13:26:38] Entity:

The Interrogator in this conversation, Transcript 6, was a non-native English speaking adult male (Chinese) who made this comment along with his 'unsure' categorisation: "I don't get enough replied information" – the judge was not fooled, effectively he did not receive enough information to make a decision as to the nature of the silent interlocutor.

On these three occasions (Transcripts 4, 5 and 6) the Judge, in each case, was unable to correctly identify the machine and, as shown, was unable to make a decision based on inadequate information from their interlocutor, so they gave their decision on the nature of the hidden entity as 'Unsure'.

#### THE RIGHT IDENTIFICATION

The concept of what is and what is not a 'right identification' is very important as far as a machine taking part in the Turing test, and the 30% pass mark, is concerned and we take a relatively strict approach in this sense. One viewpoint is that for a judge to make the 'right identification' they must correctly identify both the machine as being a machine and the hidden human as being a human (Traiger, 2000). This means that any other decision on the part of a judge would not be a 'right identification', which therefore includes cases in which either the machine is selected as a human and/or a human is selected as a machine. Also included are cases in which the judge is Unsure about either or both entities as the judge in such cases has failed to identify the machine as a machine and the human as a human – the right identification.

In our stricter interpretation however, we are only really interested in the cases in which the machine was itself not correctly identified, the judge stating either that the machine was a human or, and this is very important for this paper, that they were Unsure about it. We have in fact shown several practical examples of the latter of these two cases.

In the case of two hidden entity tests (Figure 1b), the judge also needs to make a decision concerning the human (not that the interrogator knows that the entity is a human) in parallel, sometimes selecting them as a machine, sometimes as a human and sometimes being Unsure. What we do not include, either in a general sense or as examples in this paper, are cases in which a machine was correctly identified as being a machine but where the parallel human in each case was incorrectly selected as being a machine and/or the judge gave an Unsure mark against the human, even though these would not be a right

identification. In other words we only consider cases here in which machines are incorrectly identified, either being referred to as a human or where the judge is Unsure. In each of the practical examples shown in fact it is the Unsure classification (not the right identification) which was selected by the interrogator.

### SILENCE IS GOLDEN

Turing said that in the test a machine had to try and pretend to be a man (although now/here we take that to mean human). In his 1950 paper he also pointed to the fact that at the end of 5 minutes the judge had to make a decision as to the nature of the entity. If they made the right identification and correctly identified the machine then this would effectively be a point against the machine whereas if the judge either thought that the machine was a human or if they were Unsure as to its nature then this would be a wrong identification and would be a point for the machine. The pass mark for a machine in the test was set by Turing to be 3 or more points out of every 10 (Turing, 1950).

But here we face a critical issue, what if a machine was to take the fifth amendment? The fifth amendment is an amendment to the United States Constitution establishing that, among other things, no person can be compelled to testify against himself or herself and has the right to remain silent. We extend this here to the case of 'itself' by also including the instance of machines taking part in the Turing test. In devising his test Turing specifically created the necessity for a machine to fully converse with the judge and in doing so, in the majority of cases, fail the test by giving themselves away as clearly being a machine.

But what if a machine remains completely silent during a five minute conversation? In such a case a judge receives no response to any of their questions or discussion from the hidden entity and therefore, in theory at least, cannot make the right identification and definitely say that they had been conversing with a machine. It would not be expected that a judge, under such circumstances would categorize the silent entity as being a human, although that is a possibility, the most likely case is for the judge, as we have seen in the practical examples, to give an 'Unsure' response. This of course is not a right identification and is therefore a point for the machine.

It is thus quite possible for a machine to simply remain silent to any utterances of a judge (not difficult at all to program this, in fact the computer can be removed all together) and to thereby pass the Turing test if at least 3 out of 10 judges as a result either rate the machine as being a human or indicate that they are unsure. The only thing acting against such a non-participation strategy is the fact that the machine is, in each conversation, competing against a human and if the judge is certain that the other (hidden) entity is a human then they can deduce that therefore the silent entity must be a machine. However from the practical examples it appears that interrogators frequently make a straightforward decision on the nature of a hidden machine based on that machine's performance rather than on its competitor's appearance.

It is also worth pointing out that in practice many humans are actually categorized as machines in such tests (Warwick and Shah, 2014). Therefore it is also potentially possible that a (silent) machine can be categorized as being human mainly because their hidden human competitor is categorized by the judge as being a machine. The overall possibility of passing the Turing test (and being regarded as being a thinking entity) by simply not responding at all (taking the 5<sup>th</sup>), is seen to be a major loophole in the Turing test that has hitherto not been pointed out.

## DISCUSSION

In no cases of 'silence' or failure to respond by a machine has a Judge yet classified the machine to be a human, as far as we are aware. But it is also apparent that in no cases of 'silence' or failure to respond by a machine has a Judge yet classified the machine to be a machine, i.e. the right identification, as far as we are aware. Indeed in either case it would be difficult to see how a judge could confidently make such a decision with no information to go on.

In all the four judge-machine conversations presented here, given in full in Transcripts 3-6, the judges made the decision on each occasion that they were unsure. In fact technical problems led to the machines, Cleverbot in 2014, and Brother Jerome in 2008 being unable to respond to the Judge's input. Both these machines were in fact the weakest performers, conversationally speaking, in experiments involving other machines at their respective events (see Shah and Warwick, 2010c for the best machine in 2008, Elbot and Warwick and Shah, 2015 for the best machine in 2014, Eugene Goostman).

The normal way of thinking is that the whole purpose/goal for a machine in Turing's Imitation Game is to provide the sorts of answers a human would give to the interrogators' questions. But what if the machine became truly intelligent and made a decision not to reply because it considered the human judge's question inappropriate or rude?

Turing appropriately referred to the interrogators as being judges because that is just what they are. Each judge's decision must be based on the conversational responses that an entity gives. If an entity remains completely silent then there is no way a judge can, in all honesty, make the decision that the entity is either a human or a machine. There is simply not sufficient information at their disposal. However in one judge – two interlocutor tests (Figure 1b) if it is known by the judge a priori that one entity is a human and the other is a machine then they can at least decide on the nature of a silent entity simply as being the opposite of their decision on the other entity. That said, it is known that hidden human entities are frequently identified as being machines.

In this paper we have put forward the concept that a machine can take the 5<sup>th</sup> amendment and remain silent throughout, not so much when there is technical failure, but when the machine itself chooses to remain silent. The result of this would be that a machine could potentially pass the Turing test by simply not responding at all. What we have shown however are actual transcripts in which silence on the part of the machine has resulted in the judge not making a right identification of deciding that the machine is indeed a machine.

It might be suggested that the Turing test merely needs a patch or fix of some kind in order to be operational. Perhaps if an entity is silent for longer than 15 or 20 seconds then that conversation would be immediately classified as being nul and void. The whole purpose being to attempt to avoid any (imitation) game playing by a machine. But such a decision would need to apply equally to both machine and human hidden entities and further instructions might need to be given to the hidden humans to ensure that they respond quickly enough to the last utterance from the interrogator. In a practical test there would also need to be a timer on key presses (for humans) as presently a message is only sent when the carriage return is pressed, otherwise a long sentence might be timed out. However by the same measure a machine might be programmed to enter a number of letters followed by a similar number of delete keys thereby indicating a blank space or simply a full stop as we saw in Transcript 6.

Numerous other issues arise here. Firstly instructing hidden humans to respond in a specific way appears to start to defeat the whole point of the game in that they are supposed to be themselves, humans. If they are told/programmed to behave in a mechanistic way then that will undoubtedly be easier for a machine to imitate. On top of this Turing never said anything about humans (or machines for that matter) responding in a set time. A hidden human, just like a machine, needs to be given as long as they want to reply.

Another point here can also be made with regard to counting a particular conversation as nul and void because of machine game playing. This is not an appropriate scientific response, we cannot simply ignore particular results because we do not like the look of them. In the worst case it might be appropriate to 'automatically' decide that a silent transcript results in an unsure classification. But why not give the interrogator their freedom to select, after all that is an important part of the game. Why should we take away powers of decision from a judge because of specific things that crop up in the conversation?

We are looking here at Turing's Imitation game and, particularly in its practical instantiation, the authors believe that we should stick to what Turing said about the game in order to bring it about as closely as possible to his directions. Certainly we should not unscientifically 'fiddle' the results in order to get an answer that we might prefer. Turing specifically referred to the right identification, meaning that an interrogator correctly identifies a machine as being a machine. If the interrogator does not so do then that is a point to the machine. What we have discussed in this paper is the ploy of a machine to bring about such a result by merely remaining silent.

## CONCLUSIONS

Overall it is clearly a weakness of the Turing test that it does require a machine taking part to condemn itself by its utterances, as judged subjectively by the human interrogator. But also the hidden human foil for the machine is judged similarly on their responses. Hidden humans have indeed been wrongly classified as being a machine on a number of occasions (see e.g. Warwick et al, 2013). By merely invoking the 5<sup>th</sup> amendment, the machine actually stating this in its opening rejoinder perhaps, how would a judge take this response? Would the judge think it is a smart machine trying to avoid answering the sorts of every day questions that humans engage in or would they wrongly classify it as being a human? If this was done on a sufficient number of occasions it could result in a machine 'passing the Turing test' simply by remaining silent. Of course, this does beg the question, what exactly does it mean to pass the Turing test?

Turing's game is based around a machine being discovered as being a fake human based on what it does wrong rather than what it does right. Often one particularly poor response, even if it is preceded by ten excellent responses is sufficient for an interrogator to decide that they are not talking with a human. This encourages machine developers to cause quite boring non-committal responses to often be sent, for machines to ask questions themselves or for canned sentences to be evoked. But how can a machine give itself away if it invokes the 5<sup>th</sup> amendment and simply remains silent?

If a Turing test judge is faced merely by stone wall silence in response to their utterances, how can the judge possibly tell if they are interacting (if that is what it is) with either a machine or a human? It could be a machine or a hidden human that/who has made the decision, for whatever reason, not to respond. However the lack of response could rather be caused by a technical problem with either the communication system or with the entity itself.

We opened the paper by considering that Turing introduced his imitation game as a replacement for the question "Can machines think?" (Turing, 1950). The end conclusion by many as a result of this is that if a machine passes the test then we have to regard it as a thinking machine. When/if a machine passes the test based on normal conversation then this is a nice philosophical problem which we do not wish to tinker with here. However when a machine passes the test by remaining silent this surely cannot be seen to be any indication at all of a thinking entity, otherwise we must necessarily agree that stones and rocks think. Therefore what we must finally conclude is that such a silent strategy fleshes out a serious flaw in the Turing test.

## REFERENCES

- Bourke, J., *What it means to be Human: Reflections from 1791 to the present*. Virago Press, London, UK, 2011
- Copeland, B. J., “The essential Turing—The ideas that gave birth to the computer age”, Oxford:Clarendon Press, 2004.
- Genova, J., “Turing’s Sexual Guessing Game”. *Social Epistemology*.Vol. 8, pp 313-326, 1994
- Harnad, S.,“Turing Testing and the Game of Life: Cognitive science is about designing lifelong performance capacity not short-term fooling”.LSE Impact Blog 6/10 June 10 2014 <http://blogs.lse.ac.uk/impactofsocialsciences/2014/06/10/turing-testing-and-the-game-of-life/>
- Hayes, P. and Ford, K., “Turing Test Considered Harmful”, Proc. Int. Joint Conference on Artificial Intelligence, Montreal, Vol.1, pp.972-977, 1995.
- Hodges, A. “Fair Play for Machines”, *Kybernetes*. Vol. 39, issue 3: pp. 441-448, 2010.
- Lassègue, J. (1996). What kind of Turing Test did Turing have in mind? *Tekhnema 3/ A Touch of memory/Spring*. Available here: <http://tekhnama.free.fr/3Lasseguearticle.htm>
- Moor, J. H., “The status and future of the Turing test”, In J. H. Moor (Ed.), *The Turing test – The Elusive standard of artificial intelligence* (pp. 197–214). Dordrecht, The Netherlands: Kluwer, 2003.
- Shah, H., “Deception detection and machine intelligence in practical Turing tests” PhD thesis, The University of Reading, 2011
- Shah, H., “Conversation, deception and intelligence: Turing’s question-answer game”. Chapter in (Eds). S.B. Cooper and J. van Leeuwen *Alan Turing: His Life and Impact. Part III Building a brain: intelligent Machines, practice and theory*, pp. 614-620, Oxford, UK: Elsevier, 2013
- Shah, H., and Henry, O. (2005) “Confederate Effect in Human-Machine Textual Interaction”, *Proceedings of 5th WSEAS Int. Conf. on Information Science, Communications and Applications* (WSEAS ISCA), Cancun, Mexico, ISBN: 960-8457-22-X, pp. 109-114, May 11– 14.
- Shah, H. & Warwick, K., “Testing Turing’s five-minutes, parallel-paired imitation game” *Kybernetes*, Vol 39(3), pp. 449-465, 2010a
- Shah, H., & Warwick, K., “From the buzzing in Turing’s head to machine intelligence contests”, In Proceedings of symposium for 1st towards a comprehensive intelligence test. AISB Convention, De Montfort University, UK, 29 March–1 April, 2010b.
- Shah, H., & Warwick, K., “Hidden interlocutor misidentification in practical Turing tests”, *Minds and Machines*, 20, pp. 441-454, 2010c.
- Shah, H., Warwick, K., Bland, I., Chapman, C.D., and Allen, M.J., “Turing’s Imitation Game: Role of Error-making in Intelligent Thought”, *Turing in Context II*, Brussels, 10-12 October, pp. 31-32, 2012. <http://www.computing->

[conference.ugent.be/file/14](http://conference.ugent.be/file/14)

-

presentation

available

here:

[http://www.academia.edu/1916866/Turings\\_Imitation\\_Game\\_Role\\_of\\_Error-making\\_in\\_Intelligent\\_Thought](http://www.academia.edu/1916866/Turings_Imitation_Game_Role_of_Error-making_in_Intelligent_Thought),

Shah, H., and Warwick, K. (2007). Constraining Random Dialogue in Modern Eliza, in (Schmidt C.T.A., 2007). *Computers and Philosophy, an International Conference, Proceedings*, 3-5 May 2006 Laval France,

EOARD/LIUM/AAAI/IEEE/UNESCO, London : EOARD, pp. 247-265.

Traiger, S., "Making the Right Identification in the Turing Test", *Minds and Machines*. 10, pp.561-572, 2000.

Smith, G. W. (March 27, 2015). "[Art and Artificial Intelligence](#)". ArtEnt. Retrieved March 27, 2015.

Sterrett, S. G., "Turing's Two Tests for Intelligence". In Moor, J.H. (Ed), *The Turing Test – the Elusive Standard of Artificial Intelligence* (2003) Kluwer, Dordrecht, The Netherlands, pp 79-97, 2000

Turing, A.M., "Lecture on the Automatic Computing Engine". 1947. In B. J. Copeland (Ed) *The Essential Turing: The ideas that gave birth to the computer age*. Clarendon Press: Oxford. 2004

Turing, A. M., "Computing Machinery and Intelligence", *Mind*, Vol.LIX (236), pp.433–460, 1950.

Warwick, K. and Shah, H., "Human Misidentification in Turing Tests", *Journal of Experimental and Theoretical Artificial Intelligence*, Vol.27, Issue.2, pp.123-135, 2014.

Warwick, K. and Shah, H., "Can Machines Think? A Report on Turing Test Experiments at the Royal Society", *Journal of Experimental and Theoretical Artificial Intelligence*, Accepted/to Appear, 2015.

Warwick, K., Shah, H. and Moor, J.H., "Some Implications of a Sample of Practical Turing Tests", *Minds and Machines*, 23, pp.163-177, 2013.