

# Intelligent agents: Conversations from human-agent imitation games

Warwick, K. and Shah, H.

Published PDF deposited in [Curve](#) November 2015

**Original citation:**

Warwick, K. and Shah, H. (2015) 'Intelligent agents: Conversations from human-agent imitation games' in van den Herik J.,Filipe J.,Loiseau S.,Filipe J.,Duval B. (Ed). ICAART 2015 - 7th International Conference on Agents and Artificial Intelligence, Proceedings (pp: 261-268). SciTePress

SciTePress

**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.**

**CURVE is the Institutional Repository for Coventry University**

<http://curve.coventry.ac.uk/open>

## **Intelligent Agents: *Conversations from Human-Agent Imitation Games***

Kevin Warwick<sup>1</sup>, Huma Shah<sup>2</sup>

<sup>1</sup>Deputy Vice Chancellor-Research, Alan Berry Building, Priory Street, Coventry University, CV1 5FB, UK

<sup>2</sup>Faculty of Engineering & Computing, Coventry University, CV1 5FB, UK

{<sup>1</sup>aa9839, <sup>2</sup>ab7778}@coventry.ac.uk

**Keywords:** Artificial Intelligence, Conversation, Imitation Game, Intelligent Agents, Linguistic Devices.

**Abstract:** What do humans say/ask beyond initial greetings? Are humans always the best at conversation? How easy is it to distinguish an intelligent human from an 'intelligent agent' just from their responses to unrestricted questions during a conversation? This paper presents an insight into the nature of human communications, including behaviours and interactions, from a type of interaction - *stranger-to-stranger* discourse realised from implementing Turing's question-answer imitation games at Bletchley Park UK in 2012 as part of the Turing centenary commemorations. The authors contend that the effects of lying, misunderstanding, humour and lack of shared knowledge during human-machine and human-human interactions can provide an impetus to building better conversational agents increasingly deployed as virtual customer service agents. Applying the findings could improve human-robot interaction, for example as conversational companions for the elderly or unwell. But do we always want these agents to talk like humans do? Suggestions to advance intelligent agent conversation are provided.

## 1 INTRODUCTION

Learning from human-machine and human-human text-only interactions realised from practical imitation games can assist in designing improved conversation in intelligent agents. Such agents are already in use to provide assistance with transactions in e-commerce, for entertainment purposes, and in personalised learning in education. More companies seek to virtualise their call centres with artificial customer service agents augmenting their websites for enhanced customer experience. Text-based dialogue systems are not a gimmick; they are deployed by companies such as Ikea: the Anna virtual customer agent on their website (Shah & Pavlika, 2005). In this paper we present actual discourse from *stranger-to-stranger* scenarios that provide an insight into what people ask/say to each other beyond an initial greeting. The authors contend the corpus of question-answer sequences from Turing imitation game experiments are a useful tool to innovating systems improving human-machine relations through ‘talk’.

In this paper the authors take transcripts from practical Turing tests implementing five-minute interactions at Bletchley Park, UK in 2012. We explore how humans use linguistic devices such as humour, avoidance tactics, and ‘white lies’ when interacting with strangers. Along with a machine, attempting to cause a human to make a wrong identification, two categories of humans participate in a Turing test: a) an Interrogator seeking to determine human from machine, and b) a hidden human acting as a foil for the machine. We present a series of 5-min conversations, including human-human interactions in which the interrogators wrongly categorised hidden humans as *machine*. What was it that persuaded the human interrogator that they were not interacting with another human, and did this detract from the engagement? Assumptions about what other people know are one of the features in humans misunderstanding each other.

## 2 LIVE IMITATION GAMES

The transcripts between judges and hidden entities presented here are taken from two types of tests: *simultaneous comparison* in which a judge interrogates two hidden entities in parallel (one being a conversation agent, the other a hidden human), and *viva voce* (Fig 1), where a judge questions one hidden entity at a time (Shah, 2011).

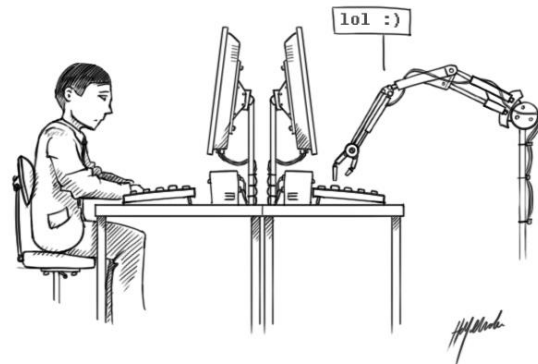


Figure 1

Traditional Turing test: Judge interrogating a machine

The imitation game, based on Alan Turing’s ideas to examine thinking and to better understand how humans think (Shah, 2013; Turing, 1952; Turing, 1950), actually involves human interrogators attempting to ascertain the nature of hidden entities with whom they are communicating. Turing raised the idea of a *first impressions* 5-minutes unrestricted questioning period after which an interrogator is tasked with making the ‘right identification’ (1950). The interrogator is required to identify the nature of their hidden discourse partner: human or machine.

When considering the game in further depth, one is quickly faced with numerous intriguing questions regarding human and machine communication and behaviour. When comparing an artificial agent’s ability to communicate with a human interrogator in natural language one immediately has to consider just who they are communicating with and the fallibility, biases and preconceptions of that person. One must also take into account important aspects of human nature such as lying, misunderstanding, unshared knowledge and humour, never mind stupidity. All important linguistic aspects that an artificial agent would need to master in order to service a call centre keeping the customer loyal, happy and trusting the brand/company.

The conversations presented here were collected as a result of five-minute long question-answer tests with human judges and hidden entities – a ‘fair play’ requirement of Turing in order that the artificial agent was not judged on beauty or tone of voice. We are fully aware that there are those who debate what exactly Turing meant: Hayes and Ford (1995)

question its purpose and feel it a distraction from successful AI. Others argue over suitable timing and number of questions in a test (see Shah & Warwick, 2010a), and there are those who take issue with the imitation game itself and believe it to be “a joke” (Marvin Minsky, 2013). The authors do not seek to respond to the controversy, rather they point to the usefulness of the corpora generated from assembling human-artificial agent conversations.

In this paper we present a number of transcripts taken from a day of practical Turing tests, which were held under strict conditions with many external viewers at Bletchley Park, UK on 23<sup>rd</sup> June in 2012. The date marked the 100<sup>th</sup> anniversary of Turing’s birth and the venue was the centre of WWII codebreaking where Turing led a team to crack the German naval enigma machine cypher (Hodges, 1992). In the experiment of 180 Turing tests five different conversational agents (chatbots) took part in a day of tests involving 30 different interrogator-judges and 25 hidden humans. The machines were compared and scored in terms of their conversational ability. One important aspect of this paper is what can be learnt from the operational performance of the human judges and specifically how they interacted in conversation with hidden entities.

Acting as foils for the machines, the hidden humans are, ‘by definition, human’, but as has been previously described (Shah & Warwick, 2010b; Warwick & Shah, 2014b) they can be misidentified on occasion (as female rather than male and vice versa) or incorrectly classified as machine – instance of the *confederate effect* (Shah, et al., 2012; Shah & Henry, 2005). This paper is concerned more with the ‘human’ involved in practical Turing tests focused on *what* they say and *how* they say it. Along a spectrum, some humans are loquacious others tend towards introversion and many fall in between. Accordingly, an attribution of *humanness* by a human interrogator to a hidden interlocutor in a practical imitation game is dependent on the judge’s own values of what constitutes humanlike conversation. Good performance of machines, with numerous examples, is discussed elsewhere (Warwick & Shah, 2013), although we do give an example here for comparative purposes.

In the sections that follow, we look at different examples of practical imitation games and attempt to cover a wide range of problem areas, which the game between intelligent human and artificial agent highlights. The transcripts considered in this paper appear exactly as they occurred; we have not altered the sequence. Once an utterance was output, it was not possible for the interrogator or hidden entity to

alter it in any way. The timings shown are accurate, actual timings on the day (UK time). Any spelling mistakes or other grammatical errors were exactly as they occurred; they are not due to editorial errors. In the transcripts, the interviewer/judge is always denoted as ‘Judge’ whereas the hidden interlocutors, machine or human, are denoted as ‘Entity’.

## 2.1 Natural Conversation

In this sample interaction we give an idea of a typical discourse exchange over a total length of 5 minutes. The exchanges give an indication of the sort of responses from hidden interlocutors that an interrogator uses to determine human or machine.

Transcript 1:

[15:44:55] Remote: hi  
[15:44:58] Judge: hi  
[15:45:06] Entity: how's your day been so far?  
[15:45:12] Judge: very interesting  
[15:45:15] Entity: why?  
[15:45:27] Judge: i have been quite busy  
[15:45:31] Entity: doing what?  
[15:45:47] Judge: I was working this morning  
[15:45:53] Entity: oh, me too  
[15:46:07] Judge: oh  
[15:46:30] Entity: i was giving a talk at the Apple Store in Regent Street. Have you been?  
[15:46:51] Judge: I think so- though I know the one in Covent Garden better.  
[15:47:04] Entity: what do you like about that one?  
[15:47:23] Judge: Apple staff are always really helpful- and I love Apple  
[15:47:40] Entity: yes they are. the stores are more about playing than buying don't you think?  
[15:48:25] Judge: Yes most of the time- that is one of the great things about Apple  
[15:48:54] Entity: what's your favourite Apple product?  
[15:49:04] Judge: My little ipad nano  
[15:49:22] Entity: cool. what colour is i?  
[15:49:30] Judge: orange

A lot of banter occurred in this discourse with a number of topical issues were covered. At the end of the conversation the interrogator quite rightly decided that they had been communicating with a hidden human. However until the topic of ‘Apple’ was mentioned – about half way through the discourse – the interaction was fairly bland with little substance. Some conversations do in fact end this way after the 5-minute total, which makes it very difficult for an interrogator to make a right decision, as there is little to go on. Clearly a ‘good’ interrogator is one who will use the time effectively asking questions that draw emotional responses. Challenges with arithmetic questions lead both

human and machine to feign incapacity (see Shah & Warwick, 2010b). Importantly in a Turing test merely asking a set of prepared questions is not as successful as facilitating a conversation of emotional depth.

## 2.2 Lying

Lying is a part of human nature and therefore has a role to play when it comes to the Turing test. The machine's goal is to cause the human interrogator to make a wrong identification, to mislead the interrogator into believing they interacted with another human. The role of the hidden human in comparison is to be themselves, *human*, whilst not giving away personal details, as this might aid the interrogator. Apart from that the human foil can fabricate, if that is their preferred response strategy. Lying can take on many different forms from a white lie, to an unintentional lie, to a complete untruth. What we give here are a couple of examples. What we are interested in is the effect of a lie on the decision taken by the interrogator. See Warwick & Shah (2014a) for an in-depth analysis of these and many more transcripts.

Transcript 2:

[12:43:23] Judge: Why hello there!

[12:43:41] Entity: Why hello to you too!

[12:44:51] Judge: How are you feeling on this fine day?

[12:45:12] Entity: To be quite honest a little rejected, I thought you were never going to reply :(

[12:45:42] Judge: Oh, I'm very sorry, it will not happen again.

[12:45:49] Entity: It just did!

[12:46:46] Judge: Oh, I lied then.

[12:47:12] Entity: That's a great shame indeed.

[12:47:34] Judge: It is. Are you following the Euro 2012's at the moment?

[12:47:55] Entity: Yeah quite closely actually. I am Cristiano Ronaldo.

The main issue with this transcript occurs in the last couple of lines. The Euro 2012 refers to the European nations' football competition which was being held at exactly the same time as the Bletchley Park Turing event. Many English people were watching the matches on television. Cristiano Ronaldo is a famous Portuguese footballer. The last line reply, merely agreeing with the judge's specific question appears to have been sufficient to categorise the entity, in the judge's opinion, as being a human. This was probably sealed by the 'humorous' comment with the entity claiming to be Cristiano Ronaldo. The interrogator may well have seen the humorous lie as supporting evidence of the entity being human. In this case the white lie had no

negative effect. This judge correctly identified this hidden interlocutor as human.

Transcript 3:

[15:44:56] Judge: hi there

[15:45:06] Entity: Hello.

[15:45:15] Judge: what is your name?

[15:45:35] Entity: I do not have a name.

[15:46:10] Judge: do you go to school or work for a living?

[15:46:24] Entity: I don't recall doing any of those.

[15:47:11] Judge: what do you have planned for this weekend?

[15:47:43] Entity: Nothing. What about you?

[15:48:22] Judge: I'm at bletchley park today and sevenoaks tomorrow, have you been to either before?

[15:49:15] Entity: I haven't been to either.

[15:49:41] Judge: can you do me a mathematical question please, what is the square root of 7

This conversation appeared hard work for the judge. Evasiveness of the hidden interlocutor in this test blocked a flowing conversation, even though they were the human. The lie by the hidden human that they had not been to Bletchley Park is clearly incorrect because they were in attendance at this venue, location for the event. However, the hidden human may have misunderstood the question to mean *had they previously visited*. If so, and they had not been there before, then they could have felt that they were telling the truth. Similarly stating that they do not have a name was a rather strange statement to make, taking security of their personal ID too far. In this case the judge's decision that the hidden entity was a machine seems defensible. The hidden human's responses were generally tame and gave nothing away. So it may be a case here of the hidden human genuinely believing they were telling the truth, when in fact they were not, possibly due to a misunderstanding. Whatever the case, the judge made an incorrect classification as a result of the hidden interlocutor's responses.

## 2.3 Misidentification

In this section we include two cases in which a misidentification has occurred. The second of these could be regarded as a good outcome in that it involved a machine being incorrectly classified as a human. The first case however involved a human misidentification (see Warwick et al., 2013).

Transcript 4:

[10:58:45] Judge: Hi there

[10:58:55] Entity: Hi, how are you?

[10:59:02] Judge: I'm good.

[10:59:21] Entity: what, that's your name, like in the Good Life?  
 [10:59:36] Judge: What do you mean by 'Good Life'?  
 [10:59:47] Entity: Tom and Barbara  
 [11:00:02] Judge: What's that?  
 [11:00:31] Entity: A couple who left the rat race  
 [11:01:03] Judge: Oh. Is that your names? Wonderful to have left the rat race. How long ago was that?  
 [11:01:28] Entity: Not me, I'm still in it. But it's nice to dream  
 [11:02:02] Judge: Yeah - would be good. I 'semi' left it. It's much less stress  
 [11:02:31] Entity: I'm still very young so I'm looking forward to a lot of things  
 [11:03:11] Judge: What kind of things?  
 [11:03:34] Entity: Life in general I guess, different emotions and feelings that it brings.

The transcript contains reference to Tom and Barbara Good who appeared in a UK TV programme entitled 'The Good Life' (Good Neighbours in the USA) in which they left the rat race to grow their own vegetables and live off the land. The discourse indicates that the judge's knowledge may not have extended to this programme. They did not appear to 'understand' the utterances of the hidden entity. Possibly as a consequence the interrogator classified the hidden human as definitely being a machine. In fact this particular human, who happens to be the first named author of this paper, was not classified as being human by any of the six interrogators involved, a singular feat which is discussed much further in Warwick et al. (2013).

Transcript 5:

[12:43:37] Judge: Hi, I'm Cathy.  
 [12:43:43] Entity: Cathy! Nice to meet you! You're a woman, if I'm not mistaken.  
 [12:44:20] Judge: Right. What is your name?  
 [12:44:29] Entity: People call me Eugene. Or "Zhenya". It is a short name for Eugene in Russia and Ukraine.  
 [12:45:17] Judge: Nice name. Are you from either area?  
 [12:45:41] Entity: Agree :-) Oh, what a fruitful conversation;-)  
 [12:45:52] Judge: Do you really think so?  
 [12:46:04] Entity: No, your guess was wrong. I don't think so. By the way, I still don't know your speciality - or, possibly, I've missed it?  
 [12:46:56] Judge: What speciality? What's yours?  
 [12:47:01] Entity: Huh? And I forgot to ask you where you are from...  
 [12:47:51] Judge: The universe and you?  
 [12:48:01] Entity: Is it your hobby to ask little poor Jewish boys such tricky questions they don't have an answer for?

At the end of this discourse the interrogator incorrectly identified the machine involved as

definitely being a human. We can see here its successful ploy. The hidden agent did not answer the question posed; it changed the subject by asking another question. The third response of the machine does seem a little off in that it completely ignored the question about origin. The final line is a good example of responding to a question in an argumentative tone. In the following section we look at further examples of machines performing well.

## 2.4 Good Machines

Here we consider two examples in which machines fooled the interrogators involved to misclassify them as human. This is in essence what the Turing test is all about and hence such transcripts are significant in assessing just what aspects of human nature, in conversation at least, appear to be successfully simulated by the machine. This is discussed in much more depth in (Warwick & Shah, 2013).

Transcript 6:

[10:41:48] Judge: Hello. How are you?  
 [10:41:53] Entity: Hey.  
 [10:42:16] Judge: How do you like Bletchley Park?  
 [10:42:20] Entity: lol.  
 [10:42:39] Judge: Are you from England?  
 [10:42:47] Entity: They have Wi-Fi here in the pub.  
 [10:43:31] Judge: Which pub?  
 [10:43:38] Entity: I'm just down the pub.  
 [10:44:03] Judge: Have you ever been in a Turing Test before?  
 [10:44:08] Entity: Will this be the 5 minute argument, or were you thinking of going for the full half hour.  
 [10:44:55] Judge: Very funny. You sound suspiciously human. Do you like the Beatles?  
 [10:45:02] Entity: I'd like to get the next Dread the Fear tape.  
 [10:45:20] Judge: What is Dread the Fear?  
 [10:45:24] Entity: Dread the fear has that Steve Henderson guy in it.  
 [10:46:08] Judge: What sort of music is that? Or is it comedy?

This conversation did not start well for the machine in this test. In their third response the mention of both wi-fi and pub seemed out of context. But it got back on track with the Monty Python humorous link between 5 minute Turing tests and a comedy sketch about the Argument clinic: *Is this a 5 minute argument or the full half hour?*. Then, in response to the Beatles question a direct answer is not given but the music link remains as the conversation shifted.

The hidden entity appeared to show in-depth knowledge of a pop group named Dread the Fear, which probably convinced the judge unaware of the

music. The machine developer is American so this group and Steve Henderson may be parochial to the developer's location. Importantly the entity gave a convincing and plausible response.

Transcript 7:

[12:23:19] Judge: What is black and white and read all over?

[12:23:32] Entity: I don't have true sense of color perception I am only able to see the RGB values.

[12:24:08] Judge: That's ok. Do you like jokes?

[12:24:19] Entity: Please tell me your inclination to a good sense of humor first.

[12:24:40] Judge: A person who can laugh at one's self.

[12:24:46] Entity: Would you like to hear a joke?

[12:24:54] Judge: Yes, please.

[12:25:21] Entity: This joke has been making the rounds in robotic circles: A robot goes to the doctor and says: "My left arm is stuck. Give me a new arm." The doctor answers: "Your arm is fine. We'll have to replace the rest of you."

[12:26:06] Judge: Ha! Ha! Why did the chicken cross the road?

[12:26:22] Entity: The chicken wandered into the road because it liked the shiny colors of the cars and trucks whizzing by.

[12:27:39] Judge: You're right. It didn't care if it got to the otherside.

[12:27:50] Entity: By the way, did you ever actually see a chicken crossing the road?

[12:28:03] Judge: Only one that didn't make it.

Here humour is applied directly by the machine, curiously with a machine type reference. Nevertheless the interrogator was certain, at the end of the 5 minutes that they had been conversing with a human. In his 1950 paper Turing referred to the *argument due to disability* some humans might use as a defence against machines being intelligent – humour being one of those aspects.

The first line uttered by the entity is a little strange, but their third line is masterly. In response to the judge's comment about laughing, the entity takes control of the conversation by offering to tell a joke. Later when the judge mentions a chicken, the entity correctly follows up with a suitable response and once again takes control of the conversation by asking the judge a pertinent question. So 'understanding' what the key issue is on a couple of occasions and actually taking control of the conversation were sufficient here for the deception to occur. In this case the machine successfully gave the illusion of understanding the judge's utterances. Such understanding has been indicated on many occasions as being a basic prerequisite for intelligence, e.g. Penrose (1994), and Searle (1997).

## 2.5 Assumptions of Knowledge

It is apparent in the Turing test that very often an interrogator will resort to asking direct factual questions even though there is little to be gained by such a ploy. As Turing (1950) indicated mathematical questions are a complete waste of time, however the same conclusion is apparent in general. Consider the following example transcript:

Transcript 8:

[15:21:30] Judge: Hi, did you know the world cup is on this week?

[15:21:44] Entity: nopes.

[15:21:55] Judge: Do you like football? What team do you support?

[15:22:08] Entity: i dont watch football

[15:22:14] Judge: What sport do you like?

[15:22:33] Entity: i bit of tennis, cricket, and bits of rest

[15:22:42] Judge: I see, do you have an XBox?

[15:22:46] Entity: nopes

[15:23:04] Judge: nopes? Do you not like computer games?

[15:23:22] Entity: i do. but not on Xbox

[15:23:47] Judge: I see, can you tell me what a limerick is?

[15:23:55] Entity: no idea

[15:24:09] Judge: Who is your favourite actor? Mine is David Niven.

[15:24:24] Entity: i am not much of a movie freak either!

[15:24:43] Judge: you must have a favourite actor. Name any actor in any film you like?

[15:25:22] Entity: well in films i like Pirates of the Carrebean series, Harry Potter

[15:25:34] Judge: What did I say was on this week?

[15:25:50] Entity: football world cup

[15:26:02] Judge: lovely old job. What school do you go to?

[15:26:17] Entity: systems engineering

[15:26:26] Judge: is that a school?

Here we see a variety of questions based on the interrogator searching for specific knowledge, possibly even looking for a particular answer. It is unclear what the expected conclusions were however. Questions about a limerick and favourite movie actor in particular were responded to with a null reply, hence very little can be concluded from them. In the final lines we see an interesting situation in which the judge is either not apparently aware of the systems engineering school or is thinking in terms of a different type of school. Despite this the judge correctly identified the hidden entity as human, but the gender blur effect occurred: the human was male but wrongly classified as female. The reference to the world cup by the judge as an opening gambit was in fact incorrect it was

Euro 2012. Did the judge mistake or used that opening lie to test the interlocutor?

Transcript 9:

[11:13:15] Judge: pastasciutta

[11:13:22] Entity: what?

[11:13:44] Judge: anything familiar in this word ?

[11:14:18] Entity: I cannot see any sexual aspect at all

[11:14:42] Judge: nothing to do with sex. try again

[11:14:49] Entity: why?

[11:15:14] Judge: just asking you to try again

[11:15:23] Entity: I am bored with it

[11:16:04] Judge: c'mon a little effort btw are you in my same time zone ?

[11:16:24] Entity: I believe that the world should just have one time.

[11:17:27] Judge: why ?

[11:17:58] Entity: seems sensible, easier for all. Better with computer

There are two issues early on in this transcript. To start with the judge throws in the word “pastasciutta” as a knowledge test for the entity. The judge assumed the hidden entity could make out from the letters *p a s t a* at the beginning of the word that it refers to a sort of pasta. Secondly the entity made a response in terms of the use of the word “familiar” in a “sexual” sense. The judge appeared to be unaware of such a meaning (Fowler & Fowler, 1995). So here we firstly see a test of specific knowledge by the judge but also an apparent lack of other specific knowledge by the same judge. As a result, the judge misidentified the human entity as being definitely a machine.

### 3 DISCUSSION

This paper is not concerned with increasing the philosophical mileage of Turing’s imitation game, nor with the importance of ‘understanding’ and the critical role it plays in intelligence, because a lot of this is in the ‘mind of the understander’. Penrose (1995) statement, that “intelligence requires understanding” does not explain the mountain of misunderstandings that occur regularly between humans who know each other, let alone among strangers, and between adults and teens, parents and off-spring, natives and non-native speakers, and of course between males and females. We do not have the time to go into analysing these and many other features of human communication here. We do feel they will need to be considered in developing intelligent agents to interact with humans and to engage each other. For example robot companions will need to communicate with medical robots about

the condition of an elderly human in their care. The Turing test can be seen to play an important role in this discussion as we can conclude from some of the transcripts presented here, that in terms of conversational appearance at least: there are some intelligent humans who appear less than bright whereas there are some machines that clearly appear smart.

It can be seen from the examples given that some judges *could* be more susceptible to deception. Others have a biased perspective on ‘humanlike conversation’. This may have led judges to misclassify hidden interlocutors, even though they were tasked with initiating conversations. Judges were given the possibility of asking or discussing whatever they wanted: the conversations were ‘unrestricted’. The ‘hidden humans’ were asked not to behave like machines and to protect their identity. However, each hidden human interpreted that instruction to ‘foils for the machines’ in their own humanlike way.

Not all the invited machines were designed to imitate humans. Elbot, for example, from Artificial Solutions has a robot personality. However, all are designed to mimic human conversation and avoid correctly answering mathematical questions, as Turing had suggested. Essentially the machines are merely trying to respond in the sort of way that a human might.

Whatever the standing of the Turing test in the reader’s mind, what we hope is evident from the transcripts presented in this paper is that it is certainly not a trivial, simple exercise. Indeed it is a surprising indication of how humans communicate and how the human judges might be easily fooled based on their assumptions and individual ideas about intelligence. Insights can lead to improved design of intelligent agents, to make their conversation more humanlike and build trust between the natural and the artificial conversation agent.

### 4 CONCLUSIONS

How humans talk in *stranger to stranger* situations suggest general techniques for successful human-intelligent agent interaction, in e-commerce for example. We suggest that intelligent agents ask more, not just to improve their conversational ability, but to *understand* the human user. We recommend that developers

- a) Do not assume knowledge held by human interlocutors



- b) Appreciate that humans cannot always formulate their enquiry clearly
- c) Develop the Intelligent Agent to probe further and ask more questions encouraging human interlocutors to clarify their needs
- d) Be prepared for mischievous users who will lie to confuse the intelligent agent.

Lastly, the authors are continuing their Turing test work, following the third event at The Royal Society London in June 2014. The results from that experiment are being analysed and will be sent for peer review.

## ACKNOWLEDGEMENT

Harjit Mehroke for Figure 1. Bletchley Park, UK, the judges and hidden humans, and the developers of the machines that took part in the 2012 experiment.

## REFERENCES

- Fowler, H., & Fowler, F. (Eds.). (1995). *The Concise Oxford Dictionary of Current English* (9th ed., p. 486). Oxford: Clarendon Press.
- Hayes, P. and Ford, K. 1995. Turing Test Considered Harmful. *Proceedings of the Fourteenth International Joint Conference on Artificial Intelligence*. Vol. 1. Montreal, August 20-25: pp. 972-7.
- Hodges, A. 1992. *Alan Turing: the Enigma*. Vintage Books, London.
- Minsky, M. 2013. Singularity 1 on 1: The Turing test is a joke! Youtube video interview by Nikola Danaylov: <https://www.youtube.com/watch?v=3PdxQbOvAII>.
- Penrose, R. (1994). *Shadows of the Mind: A Search for the Missing Science of Consciousness*. Oxford University Press, UK
- Searle, J.R. 1997. *The Mystery of Consciousness*. The New York Review of Books. NY, US
- Shah, H. 2013. Conversation, Deception and Intelligence: Turing's Question-Answer Game. In S.B. Cooper & J van Leeuwen (Eds) *Alan Turing: his life and impact*. Elsevier: Oxford, UK: pp. 614-620
- Shah, H., Warwick, K., Bland, I.M., Chapman, C.D. and Allen, M. 2012. Turing's Imitation Game: Role of Error-making in Intelligent Thought. *Turing in Context II*, Brussels, 10 October
- Shah, H. 2011. Turing's Misunderstood Imitation Game and IBM's Watson Success. Keynote in *2<sup>nd</sup> Towards a Comprehensive Intelligence test (TCIT) symposium at AISB 2011*, University of York, 5 April
- Shah, H., and Henry, O. 2005. Confederate Effect in human-machine textual interaction. *5<sup>th</sup> WSEAS Information Science and Communications (ISCA)* conference, Cancun, Mexico, May 11-14
- Shah, H., and Pavlika, V. 2005. Text-based Dialogical E-Query Systems: Gimmick or Convenience?. In *Proceedings of 10th International Conference on Speech and Computers (SPECOM)*, University of Patras, Greece, October 17-19, pp. 425-428
- Shah, H., and Warwick, K. 2010b. Hidden Interlocutor Misidentification in Practical Turing Tests. *Minds and Machines*. Vol. 20 (3), August: pp 441-454
- Shah, H., and Warwick, K. 2010a. Testing Turing's parallel-paired imitation game. *Kybernetes*, Vol. 39 (3), pp. 449-465
- Turing, A.M. 1952. With R. Braithwaite, G. Jefferson, and M. Newman BBC radio debate on 'Can Automatic Calculating Machines be said to Think?', in S.B. Cooper and J. van Leeuwen (Eds). *Alan Turing: His Work and Impact*. Elsevier: Oxford, UK, 2013, pp 667-676
- Turing, A.M. 1950. Computing Machinery and Intelligence. *MIND*, Vol 59 (236), pp. 433-460
- Warwick, K. and Shah, H. 2014b. Human Misidentification in Turing tests. *Journal of Experimental and Theoretical Artificial Intelligence (JETAI)*. DOI: 10.1080/0952813X.2014.921734
- Warwick, K. and Shah, H. 2014a. Effects of Lying in Practical Turing tests. *AI & Society*. DOI: 10.1007/s00146-013-0534-3
- Warwick, K., and Shah, H. 2013. Good Machine Performance in Turing's Imitation Game. *IEEE Transactions in Computational Intelligence and AI in Games*. DOI: [10.1109/TCIAIG.2013.2283538](https://doi.org/10.1109/TCIAIG.2013.2283538)
- Warwick, K., Shah, H. & Moor, J. 2013. Some Implications of a Sample of Turing tests. *Minds and Machines*, Vol. 23, issue 2, pp. 163-177