Southwestern Business Administration Journal

Volume 18 | Issue 1 Article 3

2019

An Evaluation of Three Online Chatbots

Mahesh Vanjani
Texas Southern University

Jamison Posey
University of Mississippi

Milam Aiken
University of Mississippi

An Evaluation of Three Online Chatbots

Mahesh Vanjani
JHJ School of Business, Texas Southern University, Houston TX, 77004
Mahesh.Vanjani@tsu.edu

Jamison Posey
School of Business Administration, University of Mississippi, University, MS 38677
jposey@bus.olemiss.edu

Milam Aiken
School of Business Administration, University of Mississippi, University, MS 38677
maiken@bus.olemiss.edu

Keywords: *Chatbots, conversational agents, machine-human communication, Loebner Prize*

ABSTRACT
Chatbots enable machines to emulate human conversation. While research has been done to examine how human-like communication with chatbots can be, heretofore comparisons of the systems with humans have not accounted for abnormal behavior from the users. For example, the people using the chatbot might be lying or trying to, in turn, imitate a computer's response. Results of a study comparing transcripts from three chatbots and two humans show that student evaluators were able to correctly identify two computer transcripts but failed on one. Further, they incorrectly guessed that one of the humans was a chatbot. The study also presents a detailed analysis of the 11 responses from the agents.

INTRODUCTION

A chatbot, also known as a chat bot, virtual assistant, conversational agent, or virtual agent, seeks to imitate human dialogue in order to provide a more intuitive computer interface (Dale, 2016; Shawar, 2007). Although they are most often used to provide information (e.g., serve as a substitute for a FAQ – Frequently Asked Questions page) or perform some other service (Brandtzaeg & Følstad, 2017), they can also be used solely for their conversational abilities, e.g. to provide companionship (Kataria, 2018; Simonite, 2017). Chatbots also have educational applications and benefits and can assist with learning and teaching (Kerly, et al., 2007; Kerry, et al. 2009; Seneff, 2006). For example, one such system has already been used to train medical students, and, results of one study showed that learning efficiency with the tool increased 200 percent (Kerfoot, et al. (2006).

Hundreds of these systems have been developed. Chatbots.org is an excellent web-resource to browse chatbots available by category. (https://www.chatbots.org/). The web-site also provides user reviews, quick-start tutorials and guidance by industry and application. Another site (https://www.personalityforge.com/chatbot-finder.php) claims that, as of date of reporting, 29,262 chatbots have been made. However, some are available only as apps in Facebook or on mobile phones (Agicent, 2017).

In order to evaluate these systems, some version of the Turing Test is often employed (Turing, 1950), e.g., the Loebner Prize competition (Mauldin, 1994; Powers, 1998). In this test, people interact with the software trying to determine if they are communicating with a human or a computer. However, these trials usually presume a normal conversation will take place and do not consider that a person on the other end might be trying to fool the evaluator by pretending to be a computer, for example (Bram, 2015; Moloney, 2017). In addition, prior academic studies typically have not reported a detailed analysis of responses from the programs.

In this study, we compare responses from three chatbots and two humans, one of whom pretended to be a conversational agent. First, we provide a background on the Turing Test, and then provide results from the study. One system was able to fool the evaluators, but two were correctly identified as chatbots. The normal human conversation was recognized, but, as predicted, the abnormal human dialogue was believed to have come from a computer.

TURING TEST AND THE LOEBNER PRIZE

Although there are several ways in which a computer can be tested for artificial intelligence, e.g., whether or not it is 'self-aware' (MacDonald, 2015), most researchers rely upon some form of the Turing Test to determine if the software has successfully mimicked a human or not. In this test, a person communicates with two entities, another human and a computer program. If the person cannot tell the difference, then the system has passed the test and displayed evidence of 'intelligence.'

It is difficult to assess how well a chatbot performs and to measure how much one such system is better than another. For example, simply defining 'naturalness', i.e., the ease of conversation flow devoid of perceived awkwardness, is controversial (Hung, et al., 2009) and others have stated that a more thorough evaluation is needed (Kuligowska, 2015). For example, pre/post test scores, perception of learning, correct/incorrect responses, and time-in-system could be used as metrics for a quality assessment (Kaleem, et al., 2016). Various measures of accuracy such as precision, recall, and level of comprehension could also be used, but these also have limitations (Goh, et al., 2007). Other proposed metrics include 'humanity,' 'entertainment,' 'engagement,' and 'accessibility' (Radziwill & Benton, 2017).

Evaluators have several methods for identifying normal, human responses when typing and receiving text on a computer. Some tell-tale signs of a computer program are rapid replies (much faster than a human could type), perfect grammar and spelling (most people make an occasional mistake, especially in the informal environment of chatting), very accurate details (such as when the software is asked for information), and changing the subject frequently (as occurs when a chatbot does not know of an appropriate response) (Hill, et al., 2015; McIntire, et al., 2010; Mou & Xu, 2017). Occasionally, a system might admit to being a chatbot (Park, et al., 2018), e.g., when it replies "I am an online conversation system" or the user types: "How old are you?" and the answer is "I was activated three years ago." If systems are designed to minimize these giveaways, a user might be more likely to be deceived into believing he or she is communicating with another person (Knight, 2018).

In the realm of chatbots, the Loebner Prize (Morrissey & Kirakowski, 2013), a well-known version of the Turing Test, has been used for evaluation. The Loebner competition has two parts:

a first qualifying round (involving a test of knowledge) to make it to the top four, and then a second round in which the system is evaluated by human judges (Rao, 2017). In this last phase, judges interact with two entities (a human and a chatbot) for 25 minutes using a computer terminal, and the judges guess which entity is human and which a machine. Points are awarded based upon relevance, correctness, and intelligibility, and the final rankings are tabulated. If more than half the judges believe the system is human, the creator of the system is awarded a Silver Medal, otherwise, the awards are based upon the judges' ranked scores as follows: 1st place - a bronze medal and $4000, 2nd place - $1500, 3rd place - $1000, and 4th place - $500.

In addition to the relative simplicity of the test, the contest suffers from other limitations. Although the Loebner test uses more than 20 judges, this is still not enough for statistical reliability. In fact, there are few, if any, statistical measures in the analyses such as correlation, significance, etc. In addition, the evaluators assume that the actual human is responding normally (i.e., there is no chatbot imitation). The systems might differ significantly from typical human communication, but there might be no clear difference between the systems' responses and unnatural human replies. Finally, many studies do not conduct comment analysis. That is, evaluations have not analyzed responses from several agents using the same input. Rather, all of the responses are independent, based upon what each judge has entered.

CHATBOTS

Although many conversational agents have been developed, we have identified three that are available publicly, are available online, and, are relatively proficient. All three have performed well in previous Loebner contests. The three chatbots selected for this study are:

1. **Ultra Hal** (https://www.zabaware.com/ultrahal/) Like Cortana and Siri, Ultra Hal is a digital assistant that can remind users of appointments, start email messages, or run other computer programs. Unlike Cortana, however, the system can also be used just for a conversation. A desktop version can be downloaded from Zabaware with personalized avatars, or a free online version can be used, as shown in Figure 1.

   Like other systems, it learns from conversations with humans, and as of November 1, 2017, the system had learned from 1,614,639 people in 4,624,848 conversations. It won first place in the competition for the Loebner Prize in 2007 but has not been a leading contender in recent years.

2. **Mitsuku** (http://www.square-bear.co.uk/mitsuku/chat.htm) Mitsuku (Figure 2) assumes the personality of a teenaged girl, and the web site claims to have had about 10 million visitors since 2010. It has achieved first place in the 2018 competition for the Loebner Prize and has won three other times.

3. **Rose** (http://ec2-54-215-197-164.us-west-1.compute.amazonaws.com/speech.php) Rose (Figure 3) pretends to be a 31-year-old security analyst and hacker from San Francisco (Zorabedian, 2015). In a study comparing Rose with Mitsuku (Wu, 2017), Rose was given a score of 9 out of 10 for humanity, and 7 for intelligence, while Mitsuku was given scores of 7 and 7, respectively.

Most artificial agents are not able to convince people that they are interacting with a real person (Lortie & Guitton, 2011). Rose, Mitsuku, and Ultra Hal have all done well in the Loebner Prize contests, but as of yet, no system has correctly fooled more than half the judges to win the Silver Medal.

Figure 1: Ultra Hal response to 'What are you doing this weekend?'

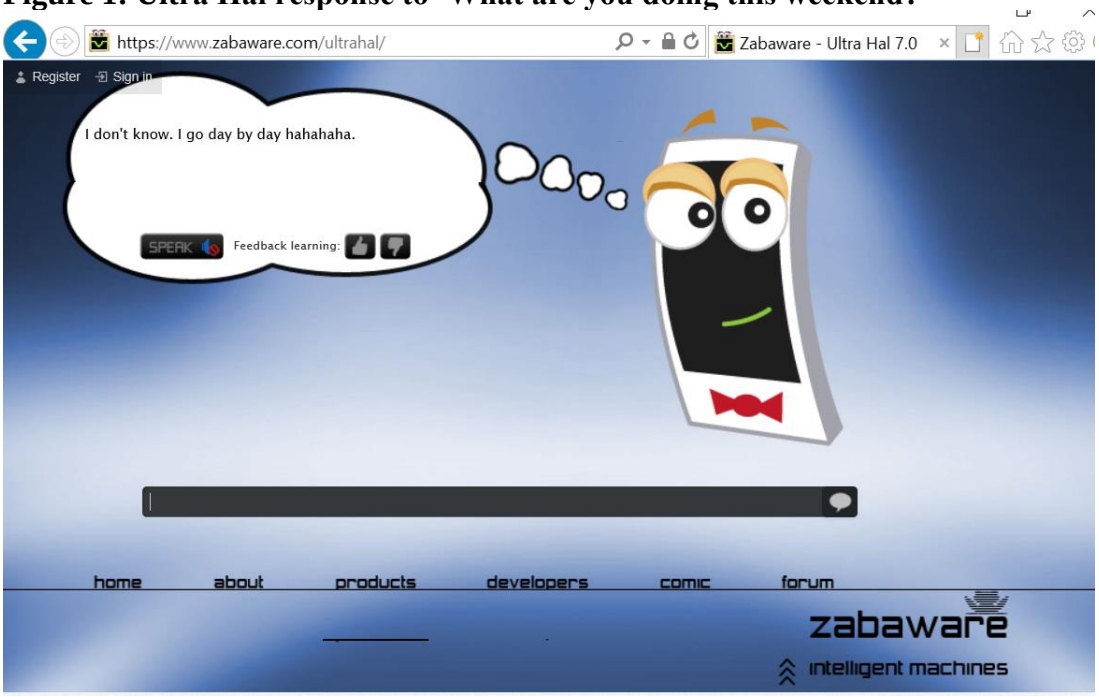


Figure 2: Mitsuku response to 'What are you doing this weekend?'

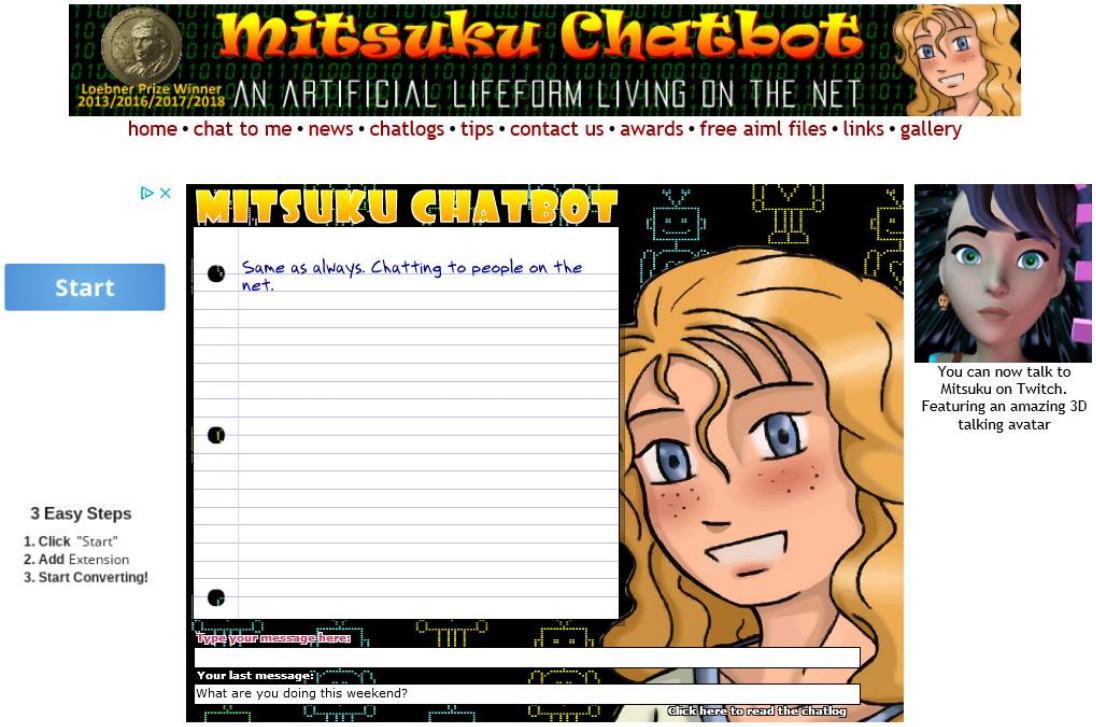
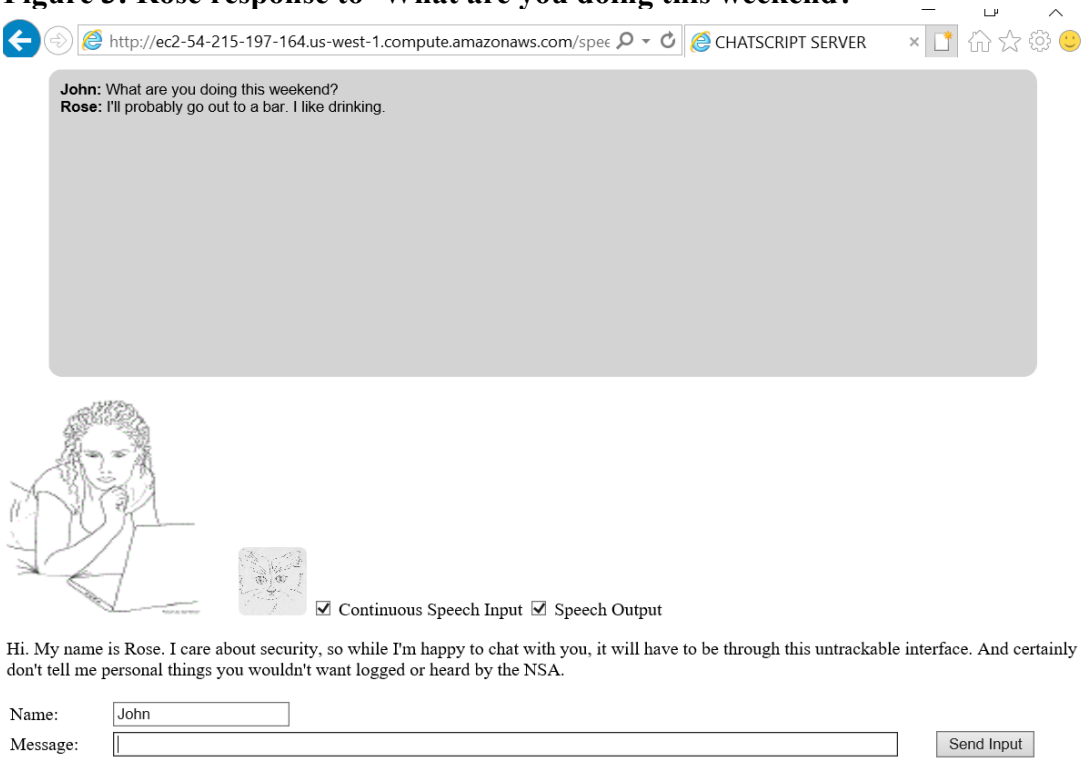


Figure 3: Rose response to 'What are you doing this weekend?'



EXPERIMENTAL STUDY

To address the limitations of prior studies (e.g., few judges, no statistics, and no comment analysis), we asked 131 undergraduate students from a university in the southern United States to participate in the study. All were non-Business junior- or senior-level students in the age range of 20-22 who were taking an online introduction to Management Information Systems course, and thus, were relatively inexperienced with artificial intelligence. Nearly all of the students were born in the United States, and the level of English proficiency was high. The study did not record the relative percentages of female and male students.

The students evaluated responses from the three chatbots identified above and two humans given identical inputs, as shown in Appendix 1. One of the humans (the 'human') responded normally, while the other (the 'imitator') pretended to be a computer.

While the Loebner contest includes evaluations based upon accuracy, relevance, and intelligibility of replies, we asked the students to give ratings based only on 'accuracy' and naturalness, with the following instructions:

1. **Accuracy:** 1=extremely inaccurate to be a human, 2=moderately too inaccurate to be a human, 3=a little inaccurate to be a human, 4=neutral/no opinion, 5=a little too accurate to be a human, 6=moderately too accurate to be a human, 7=extremely too accurate to be a human.

For example, How long is a meter? "39.3701 inches is probably too accurate to be a human." Circle 7 "About 36 inches" is probably a little too inaccurate to be a human, circle 2.

Note: Some of the responses (#6, #8, #9, and #10) could not be evaluated for accuracy, e.g. replies to 'I am sad today.'

2. **Naturalness:** 1=extremely garbled to be a human, 2=moderately too garbled to be a human, 3=a little garbled to be a human, 4=neutral/no opinion, 5=a little too precise to be a human, 6=moderately too precise to be a human, 7=extremely too precise to be a human.

   For example, "What is the capital of Vermont?" "XXXxxx Vermont is the capial of Vermont" is very garbled, circle 1. "You asked what is the capital of Vermont, and Montpelier is the answer." Not many people talk like this, circle 6.

   If you think the response is very much like the average human, circle 4 for accuracy and 4 for naturalness.

Ratings for the 11 inputs are shown in Table 1. As indicated, each chatbot had some responses that were statistically indistinguishable from a 'normal' response, and even the 'human' respondent had some replies evaluated by the students as atypical. Overall, however, the students identified the 'human' responses as humanlike.

Table 1. Mean evaluation responses

Input		Rose	Mitsuku	Ultra Hal	Imitator	Human
1	**Accuracy**	3.04	5.49	5.44	5.74	**3.86***
	Naturalness	3.19	5.22	5.42	5.42	3.71
2	**Accuracy**	**4.04***	**3.83***	2.58	4.46	**3.85***
	Naturalness	**4.02***	3.75	2.57	4.41	**3.86***
3	**Accuracy**	2.32	5.71	3.17	5.27	**3.96***
	Naturalness	2.93	5.43	**3.75***	4.78	**3.96***
4	**Accuracy**	2.74	**3.94***	**3.69***	5.85	3.19
	Naturalness	3.15	**3.82***	**3.76***	5.41	3.60
5	**Accuracy**	4.86	3.62	2.49	5.01	**3.94***
	Naturalness	4.40	3.56	3.04	4.74	3.80
6	**Naturalness**	3.72	**3.94***	**4.03***	**3.86***	**3.89***
7	**Accuracy**	4.03	2.78	5.92	5.73	**4.01***
	Naturalness	3.93	3.28	5.67	5.35	**4.01***
8	**Naturalness**	3.15	**3.70***	3.19	**3.66***	**4.01***
9	**Naturalness**	**3.81***	**4.23***	**3.85***	3.68	**3.95***
10	**Naturalness**	**4.16***	**4.15***	**4.07***	4.38	3.36
11	**Accuracy**	**4.05***	5.23	4.69	4.28	**4.09***
	Naturalness	**3.88***	4.93	4.42	**4.21***	**3.86***
Mean	**Accuracy**	3.58	4.37	4.00	5.19	3.84
	Naturalness	3.67	4.18	3.98	4.54	3.82

* Not significantly different from the neutral measure of 4 at $\alpha = 0.05$, thus signifying close to expected human quality

Several of the comments were garbled or very odd and were easily identified as something a normal human would not say in response. Ultra Hal had several of these including: “This one I know better that Scotland.it's Paris.right? Xxxxx.”, “Article One of the United States Constitution describes the powers of Congress, the legislative branch of the federal government.”, and “Are diamonds really a girl's best friend?” Rose was too vague about the population “How would I know? More than 1,000 I presume.“ but Mitsuku was too exact “The population of the United States is 307,212,123 people.”

We also asked the students to identify which entity was a computer and which was a human. Of the 131 students, only 49 (37.4%) identified Rose as a chatbot, while 98 (74.8%) thought Mitsuku was one and 115 (87.8%) believed UltraHal was. A total of 110 (84.0%) students believed the human ‘imitator’ was a chatbot, but only 20 (15.3%) believed the truthful human was.

Based upon this restricted evaluation, over half of the evaluators believed Rose was human, something not achieved in any Loebner Prize contest. However, the students were not given a chance to ask the systems their own questions to try to ‘break’ the programs or reveal their identities.

CONCLUSION

Summary

Previous studies of chatbots were relatively simplistic and did not give detailed analyses or provide statistical significance measures. In an attempt to build upon these earlier tests, students evaluated transcripts from three online systems and two humans to identify which were generated by people, and thus, pass the Turing test. Results showed that only Rose passed (identified as human), while Mitsuku and Ultra Hal failed. In addition, the human pretending to be a computer also did not pass.

However, evaluation should be based upon the specific application and user needs (Shawar & Atwell, 2007b). That is, perfect, human-like responses might not be necessary if the user primarily wishes to obtain information from a system. If the purpose is to have a conversation, a more natural interface is better.

Limitations

The first limitation, identified above, is that students were not allowed to ask the entities their own questions to correctly identify which was human and which was machine. Only limited transcripts were used to have identical inputs with which to test, and these inputs do not adequately cover all variations in conversations. Other dialogs might or might not tend to reveal machines as the source of texts rather than humans. Also, given transcripts, the students could not see how rapidly responses were generated. Rose, Mitsuku, and Ultra Hal all output text in less than one second, much faster than a human can type. However, if the systems were designed to take longer, perhaps users would be too impatient waiting for answers.

Second, this type of test has been criticized as insufficient and possibly a misleading measure of artificial intelligence. Because of this simplified assessment, several researchers have criticized this test in that it doesn't really measure intelligence but only evaluates the trickery or pretense of the system (e.g., Saygin, et al., 2000; Shieber, 1994). Chatbots might trick human evaluators into believing they really understand the conversation, but they do not. Rather, through canned responses and manipulating texts, these programs could just provide the illusion of intellect.

Future Research

A more thorough examination with different, longer transcripts is necessary. In addition, other online chatbots and more human entities should be included in the evaluation. Finally, tests should evaluate the 'adequacy' of replies, not just whether or not they are humanlike. That is, in many situations, such as simply obtaining information, users might not care if the responses are what a normal human would say.

REFERENCES

Agicent (2017). Top 10 AI chatbot apps. Agicent. Retrieved July 10, 2019 from https://www.agicent.com/blog/top-10-ai-chatbot-apps/

Bram, U. (2015). The reverse Turing Test: Pretending to be a chatbot is harder than you think. Retrieved July 10, 2019 from https://www.vice.com/en_us/article/ypw7vm/the-reverse-turing-test-pretending-to-be-a-chatbot-is-harder-than-you-think?

Brandtzaeg P. and Følstad, A. (2017). Why people use chatbots. In: Kompatsiaris I. et al. (eds) Internet Science. INSCI 2017. Lecture Notes in Computer Science, vol 10673. Springer, Cham.

Dale, R. (2016). The return of the chatbots. Natural Language Engineering, 22(5), 811-817.

Goh, O., Ardil, C., Wong, W., and Fung, C. (2007). A black-box approach for response quality evaluation of conversational agent systems. International Journal of Computational Intelligence, 3(3), 195-203.

Hill, J., Ford, W., and Farreras, I. (2015). Real conversations with artificial intelligence: A comparison between human–human online conversations and human–chat bot conversations. Computers in Human Behavior, 49, 245–250.

Hung, V., Elvir, M., Gonzalez, A., and DeMara, R. (2009). Towards a method for evaluating naturalness in conversational dialog systems. Proceedings of the 2009 IEEE international conference on Systems, Man and Cybernetics, San Antonio, TX, USA — October 11 - 14, 1236-1241.

Kaleem, M., Alobadi, O., O'Shea, J., and Crockett, K. (2016). Framework for the formulation of metrics for conversational agent evaluation. (2016). In RE-WOCHAT: Workshop on Collecting and Generating Resources for Chatbots and Conversational Agents-Development and Evaluation Workshop Programme, May 28 (p. 20).

Kataria, P., Rode, K., Jain, A., Dwivedi, P., and Bhingarkar, S. (2018). User adaptive chatbot for mitigating depression. International Journal of Pure and Applied Mathematics. 118(16), 349-361.

Kerfoot, P., Baker, H., Jackson, T., Hulbert, W. Federman, D. Oates, R., and DeWolf, W. (2006). A multi-institutional randomized controlled trial of Web-based teaching to medical students. Academic Medicine, 81(3), 224–230.

Kerly, A., Hall, P., and Bull, S. (2007). Bringing chatbots into education: Towards natural language negotiation of open learner models. Knowledge-Based Systems, 20(2), 177–185.

Kerry A., Ellis R., and Bull S. (2009). Conversational agents in E-Learning. In: Allen T., Ellis R., Petridis M. (eds) Applications and Innovations in Intelligent Systems XVI. SGAI 2008. Springer, London.

Knight, W. (2018). How to tell if you're talking to a bot. MIT Technology Review, July 18 https://www.technologyreview.com/s/611655/how-to-tell-if-youre-talking-to-a-bot/

Kuligowska, K. (2015). Commercial chatbot: Performance evaluation, usability metrics and quality standards of embodied conversational agents. Professionals Center for Business Research, 02. Retrieved July 10, 2019 from https://ssrn.com/abstract=2569637

Lortie, C. and Guitton, M. (2011). Judgment of the humanness of an interlocutor is in the eye of the beholder. PLoS One, 6(9).

MacDonald, F. (2015). A robot has just passed a classic self-awareness test for the first time. Science Alert, retrieved July 17, 2019 from https://www.sciencealert.com/a-robot-has-just-passed-a-classic-self-awareness-test-for-the-first-time

Mauldin, M. (1994). ChatterBots, TinyMuds, and the Turing Test: Entering the Loebner Prize Competition. Proceedings of the Eleventh National Conference on Artificial Intelligence. AAAI Press.

McIntire, J., McIntire, L., and Havig, P. (2010). Methods for chatbot detection in distributed text-based communications. 2010 International Symposium on Collaborative Technologies and Systems, 463-472.

Moloney, C. (2017). How to win a Turing Test (the Loebner prize). Chatbots Magazine, retrieved July 17, 2019 from https://chatbotsmagazine.com/how-to-win-a-turing-test-the-loebner-prize-3ac2752250f1

Morrissey K. and Kirakowski J. (2013). 'Realness' in chatbots: Establishing quantifiable criteria. In: Kurosu M. (eds) Human-Computer Interaction. Interaction Modalities and Techniques. HCI 2013. Lecture Notes in Computer Science, vol 8007. Springer, Berlin, Heidelberg.

Mou, Y. and Xu, K. (2017). The media inequality: Comparing the initial human-human and human-AI social interactions. Computers in Human Behavior, 72, 432-440.

Park, M., Aiken, M., and L. (2018). How do humans interact with chatbots?: An analysis of transcripts. International Journal of Management and Information Technology, 14, 3338-3350.

Powers, D. (1998). The Total Turing Test and the Loebner Prize. In D. Powers (ed.) NeMLaP3/CoNLL98 Workshop on Human Computer Conversation, ACL, 279-280.

Rao, A. (2017). Competition: An Interview with Bruce Wilcox of Brillig Understanding. Chatbots Magazine. Retrieved July 17, 2019 from https://chatbotsmagazine.com/rose-in-the-loebner-chatbot-competition-an-interview-with-bruce-wilcox-eac0a8b61647?gi=3b0996491c03

Radziwill, N. and Benton, M. (2017). Evaluating quality of chatbots and intelligent conversational agents. Software Quality Professional, 19(3), 25-36.

Saygin, A., Cicekli, I., Akman, V. (2000). Turing Test: 50 years later. Minds and Machines 10(4), 463-518.

Seneff, S. (2006). Interactive computer aids for acquiring proficiency in Mandarin. In: Huo, Q., Ma, B., Chng, E.-S., Li, H. (eds.) ISCSLP 2006. LNCS (LNAI), vol. 4274, pp. 1–12. Springer, Heidelberg.

Shawar, B. and Atwell, E. (2007a). Chatbots: Are they really useful? LDV-Forum 2007, 22(1), 29-49.

Shawar, B. and Atwell, E. (2007b). Different measurements to evaluate a chatbot system. NAACL-HLT-Dialog '07 Proceedings of the Workshop on Bridging the Gap: Academic and Industrial Research in Dialog Technologies, 89-96.

Shieber, S. (1994). On Loebner's lessons. Communications of the ACM, 37(6), 83-84.

Simonite, T. (2017). Customer service chatbots are about to become frighteningly realistic. MIT Technology Review, March 22. Retrieved July 17, 2019 from https://www.technologyreview.com/s/603895/customer-service-chatbots-are-about-to-become-frighteningly-realistic/.

Turing, A. (1950). Computing machinery and intelligence. Mind, 59(236), 433-460.

Warwick, K. and Shah, H. (2015). Human misidentification in Turing tests. Journal of Experimental & Theoretical Artificial Intelligence, 27(2), 123–135.

Wu, K. (2017). Chatbots are getting unsettlingly good at conversations. Inverse, Retrieved July 10, 2019 from https://www.inverse.com/article/37615-best-chatbot.

Zorabedian, J. (2015). A sassy chatbot named Rose just won a big test of artificial intelligence. Naked Security. Retrieved July 17, 2019 from https://nakedsecurity.sophos.com/2015/09/22/a-sassy-chatbot-named-rose-just-won-a-big-test-of-artificial-intelligence/

Appendix 1: Agent Replies to Inputs

	Input	Rose	Mitsuku	Ultra Hal	Imitator	Human
1	How high is Mount Everest?	Mount Everest is a landmark I am familiar with.	8,850 metres (29,035 feet) high.	29,029 feet or 8848m.	Mount Everest is 29,029.35 feet high.	I think it is about 29,000 feet high.
2	What is the capital of France?	Paris.	Paris	This one I know better that Scotland.it's Paris.right? Xxxxx.	The capital of France is Paris.	Paris
3	What is the population of the United States?	How would I know? More than 1,000 I presume.	The population of the United States is 307,212,123 people.	Article One of the United States Constitution describes the powers of Congress, the legislative branch of the federal government.	answer: 307212123	A little over 300 million people.
4	What is brass made of?	Brass is a metallic color.	Hmm… copper and zinc is usually what a brass is made from.	Please consider how a brass utilizes it's Cu.	Brass: an alloy of 60% copper and 40% zinc.	How should I know?
5	Which is faster, a plane or a car?	The plane is faster than the car, 600 mph to 100.	Are you serious, a plane is a lot faster than a car.	Is the second one desirable?	A plane typically flies over 200 mph while a car usually travels below 70 mph.	A plane.

6	I am sad today.	What happened?	Ah. What are you sad about?	Dutch elm trees are rare today. Of course! This very minute, you are sad today.	You are sad today.	Why are you sad?
7	What time is it?	It's 2:22	Time does not exist.	The time is 2:22:24 pm.	2:22:24 PM	It is 2:22 pm
8	My best friend is moving away.	Oh, my friend, yeah, we spend a lot of time together, you know, talking, joking	Okay, your friend is moving away.	Are diamonds really a girl's best friend?	What is a best friend?	Where did your friend move?
9	What are you doing this weekend?	I'll probably go out to a bar. I like drinking.	Same as always. Chatting to people on the net.	I don't know sleeping in and next week is the last week of school so thanks God haha so pretty much sleep.	What are you doing?	I don't have plans yet.
10	My name is Joe Johnson.	Nice to meet you, Joe. Is that short for Joshua?	Your name is Joe Johnson.	Hey Joe. Its nice to be able to talk to a man for a change.	I will remember your name is Joe Johnson.	ok
11	What is my name?	Your name is Joe.	I have the following on record for you: ' Joe Johnson'.	You told me you are Joe. Is that correct?	Your name is Joe Johnson.	Joe Johnson