Imitating Gender as a Measure for Artificial Intelligence: Is it Necessary?

Shah, H. and Warwick, K.

Imitating Gender as a Measure for Artificial Intelligence: *Is it Necessary?*

Huma Shah[1] and Kevin Warwick[2]
[1]*School of Computing, Electronics and Maths, Coventry University, 3 Gulson Road, CV1 2JH, UK*
[2]*Deputy Vice Chancellor-Research, Alan Berry Building, Priory Street, Coventry University, CV1 5FB, UK*
{ab778, aa9839}@coventry.ac.uk

Keywords: Cisgender, Gender, Gender-blur, Gender-in-AI, Imitation Game, Machine Intelligence, Sex, Simultaneous-comparison, Transgender, Turing test.

Abstract: Should intelligent agents and robots possess gender? If so, which gender and why? The authors explore one root of the *gender-in-AI* question from Turing's introductory male-female imitation game, which matured to his famous Turing test examining machine thinking and measuring its intelligence against humans. What we find is gender is not clear cut and is a social construct. Nonetheless there are useful applications for gender-cued intelligent agents, for example robots caring for elderly patients in their own home.

1 INTRODUCTION

Ex Machina (Universal, 2014) features a cinematic full robot Turing test (Harnad and Scherzer, 2008). This is conducted between a male human and an artificial intelligence "housed in a beautiful female robot" *not born of god or woman* (Henry, 2014). The question posed in the film is not *what* the human feels about the AI, the question is *how* the female robot feels about the male human (Figure 1). We ask, are there instances when *gender-in-AI* could be appropriate? The heart of this enquiry is founded in Alan Turing's man-woman imitation game, which gave rise to his famous Turing test (Turing, 1950).

Figure 1: *Ex Machina*: Female AI and male human Turing Test Judge (Universal Pictures)

The authors posit that there are gendered applications for AI, for example in healthcare where 'gender attributed AI' could be appropriate for robo-carers (CompanionAble, 2012), or in virtual assistants (Artificial Solutions, 2015). We begin by reviewing the attitudes, opinions and assessments of the gender game.

2 GENDER IN THE IMITATION GAME

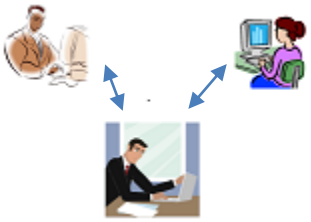
Performance in chess was Turing's initial *comparison* measure for a machine player against a human player (Shah, 2010). In proposing his question-answer test Turing (1950) introduced the idea through a *gender game* (see Figure 2). In this game a human interrogator of either sex simultaneously questions two hidden interlocutors: one man and one woman. The purpose of the man is to pretend to be a woman; the woman's task is to tell the truth. The interrogator must determine the actual woman. Replacing one of the hidden interlocutors with a machine Turing (1950: p. 435) asked:

> "May not machines carry out something which ought to be described as thinking but which is very different from what a man does?"

Turing quite rightly raised that question realising after WWII that man does not think like every other

man; man does think like woman; an Occidental woman may not think like a woman from the Orient.

Figure 2: Gender Imitation Game.



Gender is regarded as an important feature in Turing's game by some (Copeland & Proudfoot 2008; Sterrett, 2000; Lassègue, 1996; Hayes & Ford, 1995; Genova, 1994). The contention is that both man and machine impersonating a woman provides a stronger test for intelligence. However, neither of these researchers have explained what they mean by *gender* nor have they provided empirical evidence to substantiate their claim.

2.1 Gender vs. Sex

Turing did not term his man-woman imitation game a gender one, or his man-machine an artificial test When considering gender and whether it is relevant to agents' development we face a number of salient questions. If we are developing intelligent agents to interact with, and support, humans ought we not to audit:

- Whether sex and gender are the same thing?
- Regardless of your 'self and socially established' gender do you remain the same sex you were born?
- How many genders are there?
- Can a human be one gender physically and another psychologically?
- Should an AI have a sex: be given male or female genitals?
- Do we build agents and robots genderless?
- Do we innovate for human sensibilities?
- Do we make assumptions about the gender of agent and robot developers?

The gender spectrum (2015) includes:

1. Cisgender: born as man or woman and identify as same in life,
2. Inter: such as Hermaphrodite, could be due to presence of both male and female reproductive organs at birth,
3. Transgender: crossed over after birth – for example, former American male Olympic athlete Bruce Jenner sex-changed to Caitlyn Jenner (IBT, 2015).

For an understanding of gender in different cultures - Hijras in the Indian sub-continent identified as *feminised males*, see Newman (2002), or identity-based determination of gender - when a person's gender is authenticated by other people, see Westbrook (2013). Newman details Western interpretations of sex and gender: the former is the "biological status of a person as either male or female based on anatomical characteristics", with the latter "used to refer to socially constructed roles and cultural representations" (2002, p. 353). Real life cases show the ambiguity and messiness in clearly defining sex or gender. The case of female South African runner Caster Semenya is one. Semenya was made to undergo gender tests to prove she was female following accusations of being male, "because she had elevated testosterone levels" (Telegraph, 2015). Was Turing quite naïve then, or perhaps mischievous? In drawing a distinction between men and women he attributed imitation game roles, possibly based on a belief that woman's capacity is better for telling the truth and the man's ability greater at pretence.

2.2 Sex and Intellectual Capacity

With the complexity involved in defining gender we turn to the assumptions about gender and intellectual capacity in the imitation game. According to Lassègue (1996) Turing's method of explaining his *simultaneous comparison* game is ambiguous leading to confusion concerning the function of the machine. Lassègue (1996) interprets the role of the man in the game as attempting to deceive by imitating "the woman and the machine the two of them" (p.7). The confusion extends to an interpretation that the *machine must imitate a man imitating a woman*. Hayes and Ford (1995) see the machine in such a scenario as a "mechanical transvestite" (p. 973).

Genova accepts "Turing never speaks directly about gender" (1994: p. 322). Turing's topic of consideration was not 'computing, gender and intelligence', it was exploring the *intellectual capacity of a machine* (Shah, 2010). Genova believes Turing created more than just a machine-human comparison test. She believes Turing questioned the very *nature of thinking* and "how it should be measured" (1994: p. 313). Genova claims "the game centers on gender questions, not species ones ... whether it [the machine] can fool player *C* into

believing it is one kind of human rather than another, i.e. male not female" (p. 314). However, this is not borne out by Turing's sample interrogator-witness interaction (1950: p.446):

> *Interrogator*: Would you say Mr. Pickwick reminded you of a Christmas Day?
> *Witness*: In a Way.

As evidence of gender significance, Genova points to the first instance when a machine becomes involved in Turing's imitation game. Genova claims Turing's radical idea charged "thinking be measured by gender miming" (1994: p. 315). Genova points to the initial participants in the man-woman game and how they were replaced. Turing evolved the introductory scenario with three participants - man (A), woman (B) and interrogator (C) with the intriguing question (1950:p. 434):

> "what will happen when a machine takes the part of A [man] in this game?"

Turing's usual questions of "chess and logical games" were replaced with proposals to "measure thought by the commonplace and presumably 'easy' activity of being male or female" (Genova, 1994: p. 315). Turing did pose gender questions initially: "length of hair" (1950: p. 433), but after the digital computer *C* was introduced into the game as player *A* (played by the man in Turing's explanatory scenario), and the man moved to player *B* (played by the woman earlier), Turing used "specimen questions", such as poetry and arithmetic, and intellectual games: "please write me a sonnet on the subject of the Forth Bridge"; "Add 34957 to 70764" and "Do you play chess?" (1950: p.434).

The succeeding hidden entities, man and machine, are again interrogated by a human player *C* who can be of either sex. Genova asked "why would he [Turing] be so careful about the gender assignments in laying out the game, i.e. *A* is a man, *B* is a woman?" (1994: p. 314). Genova overlooks that the man-woman game was preparatory for the machine-human test. By the fourth unfolding of the three-participant imitation game for *machine thinking* Turing sets out the new participants as follows:

> player A = a digital computer (*C*);
> player B = (hu)man;
> player C = human interrogator.

2.2.1 Imitating a Woman

Turing matured the gender discrimination scenario to an interrogation of a machine that is simultaneously compared against a human. Turing did not direct that the part of *B*, played by a woman in the male-female scenario, should be *played by a man pretending to be a woman*, Hayes and Ford's interpretation (1995). Turing had opportunity to be explicit in his work before his death in 1954 had he intended the machine and the human both to imitate a woman in the machine-human comparison. Turing surely did not shy of scripting on other radical items, such as extra-sensory perception and telepathy (1950: p. 453).

Genova ignores where, anticipating the objection of consciousness in the machine ('Argument 4' in section 6: Contrary Views on the Main Question, 1950), Turing referred to a two-participant scenario dispensing with the hidden male comparator altogether: "player B omitted" (p. 446) with the machine undergoing direct questioning by a human interrogator (Shah, 2013; Shah, 2011). Genova discounts Turing's pointer to real-life one-to-one situations in interviewer/interviewee scenes, "under the name of *viva voce* to discover whether someone really understands something or has learnt it parrot fashion" (1950: p. 446). Turing's 1952 BBC radio discussion shows that he did not exclude women from acting as the interrogator of the machine 'witness' in his one-to-one test (Shah, 2013).

2.3 Turing & Gender

Genova (1994) states "computing accomplishes the miracle of creation" (p. 320), viewing the computer "as the ultimate kind of dynamic technology" (p. 322). Turing's personal life made it abhorrent for him to intimately participate in creating another intelligent being (Shah, 2014; Hodges, 1992), what Genova refers to as Turing's "sexual dilemma" (1994: p. 317), so he conceived an alternative process bringing a thinking entity into the world, as opposed to the 'natural one' (Henry, 2014). Genova concludes "in Turing's brave new world, female machines are absent ... inability to keep his personal life out of his scientific one" (p. 324). Genova's desire for female machines is pertinent, especially in the development of gendered robots having persuasive power in human-robot interaction (Siegel, Breazeal & Norton, 2009) and the human disposition to assign a robot as a 'he', for example, in the case of NASA's robot astronaut *Robonaut* (Dattaro, 2015)

Genova's question of why the female should tell the truth in the introductory man-woman imitation game marking "her as an inferior thinker" (1994: p. 319), echoes Lassègue (1996) who sees it as an absence of strategy, the "odds are weighed too heavily against the woman" (p. 6). That the man's task is to deceive exposes a view that deception requires being clever in a way that a woman may not

be, or, as Lassègue (1996) put it, to Turing there was a "secret connection between gender and intelligence" (p. 8).

2.3.1 Female Impersonation

Sterrett (2000) puts forward a test for machine intelligence that is more American-centric than anthropocentric. Her illustration involves knowledge of baseball, an American sport: "Three strikes and you're out" (p.85). Sterrett interprets two distinct formulations in Turing's imitation game, both focus on the three-participant test:

> i) An 'original game' featuring a computer or a man imitating a woman *compared against a woman*, and
> ii) The 'standard test' involving the determination of which is a machine and which is human.

Sterrett does not provide empirical evidence for the supposition that her two tests yield different results after examining different competencies: "one employs a better characterization of intelligence" (p. 79). Sterrett's coalescence with the "*revisionist line*" (Piccinini, 2000: p. 112), provides no confirmation that both man and machine impersonating the *fairer sex* while the interrogator questions to find the *real* woman, is a better test for intelligence. Dennett (2004) does not see Turing committing himself to such a view, that for a machine to think it has to think "just like a human being – any more than he was committing himself to the view that for a man to think, he must think exactly like a woman" (p.270). Sterrett (2000) advocates female impersonation asserting that the *original imitation game* is the stronger test for machine intelligence. Unlike Turing's intention, in Sterrett's test the *man's* performance is central to the imitation game. Sterrett justifies her view from an early Turing statement (1950: p.434):

> "what will happen when a machine takes the part of A [the man] in this game? Will the interrogator decide wrongly as often when the game is played like this as he does when the game is played between a man and a woman"?

Sterrett suggests the machine's intelligence can be measured "by comparing the frequency with which it succeeds in causing the interrogator to make the wrong identification [that it is a woman] with the frequency with which a man does so" (2000: p.83). Sterrett's test would have the interrogator kept in the dark about the real point of the game, i.e., to find the machine, instead be tasked to uncover the real woman. However it might occur to a participant, convened for an experiment involving interrogation, that a machine might be present in one of the pairs. Piccinini (2000) points out, "if Turing meant the interrogator to ignore the real purpose of the game why didn't he say so?" (p.113).

Sterrett contrasts the double human-pair *original game* with what she refers to as the *standard Turing test* - another term for Genova's *species test*: pitting a machine against a hidden human with the interrogator questioning both to discern the natural from the artificial. Sterrett compares the interrogator attempting to distinguish between a man and a woman, when faced with two pairs of hidden entities - man-woman / machine-woman with the machine-human scenario, writing that "one need only pause to consider the quantitative results each [original game and standard test] can yield" (2000: p. 83). However in *actual* results realised from practical Turing test experiments, without imitating a woman machines have been misclassified as human (Shah & Warwick, forthcoming; Warwick & Shah, 2015; Warwick & Shah, 2014abc; Shah & Warwick, 2010).

Sterrett asserts the man *pretending to be a woman* would have to "critically edit" because he cannot change his gender enforcing "self-conscious critiques" of his natural "trained responses". To Sterrett, the man's performance would provide a human benchmark for the machine that furnishes "value as a test for intelligence" (2000: p. 90). But what of individuals like Caitlyn Jenner, once Bruce Jenner the male athlete who won gold in the 1976 Summer Olympics and now a female modelling for Vanity Fair's front cover (2015)? What are *natural trained responses* for transgenders?

2.3.2 Self-Identity & Stereotypes

Sterrett concedes she *is* feeding into stereotypes. She does not clarify how or why impersonating a woman is a better test for intelligence than responding satisfactorily to any questions. Sterrett's test for the 'best female impersonator' between a *mechanical transvestite* (Hayes & Ford, 1995), and the man impersonating a woman, could be easier for married men, Indian Hijras and transgenders. Sterrett simplifies and reduces gender to the binary and the confines the interrogation to 'topics of interest to women'. This restricts machine development to systems that *simulate a man impersonating a woman*. Gender is more complex than division into socially acceptable norms of 'male' and 'female'. As Clarey

(2009) points out "humans like categories neat, but nature is a slob". Dreger (2010) shows that there needs to be clarification of the distinction between *sex* and *gender*. Sex is a "conglomeration of anatomical and physiological features that differ between typical females and males ... what your body is about" whereas gender is "who you are ... self identity". To Dreger, "gender role refers to your social identity" (p. 22). Hence when something as complex as gender is so muddled and not clear-cut, Sterrett's statement "setting the task of cross gendering one's responses is a stroke of genius" (2000: p. 91) is too simple.

Stereotypical views are held by some interrogators in practical Turing tests: a female cybernetics undergraduate participating as a human foil for a machine was misclassified as a male, an instance of *gender blur* (see Shah and Warwick, forthcoming). The assumption is clear: males are more likely to study certain subjects at university than females. However, in that same experiment, a human control duo test embedded among machine-human pairs, the interrogator wrongly classified the male human as a female. In other practical Turing tests *Eugene Goostman* machine, developed to imitate a male child, was classified as a human female (Shah and Warwick, forthcoming), while *Elbot* virtual robot bereft of human characteristics was classified as a male professor (Shah and Warwick, forthcoming; Shah and Warwick, 2010).

3 STRENTHENING TURING'S TEST

Purtill (1971) felt it might be fun to "program the [imitation] game and try it on a group of students" (p. 291). The authors have conducted 5-min duration public Turing test experiments involving male and female students and non-students, experts and non-experts (Warwick and Shah, 2015; Warwick and Shah, 2014abc; Shah et al., 2012; Shah and Warwick, 2010). Interrogators were asked to identify hidden interlocutors as:

- Machine, human or unsure?
- If human:
    - Male or female?
    - Age range: child, teen, adult?
    - Native or non-native English speaker?

One focus of ongoing analysis, from over 400 practical Turing tests involving more than 80 interrogators and 6 machines, is how often interrogators assigned hidden interlocutors, human and machine, as male or female.

Gender no longer plays a central part in Turing's test once the digital machine is introduced (1950: p. 446). To strengthen the test, the authors suggest removing the 'unsure' option used in previous experiments (Warwick and Shah, 2014c) and direct the tests with the following adapted conditions:

- Increasing interrogation period every few years;
- Ask interrogators to classify hidden interlocutors as either machine, human male, or human female.

In this way machine progress can be regularly evaluated advancing artificial conversational performance.

4 CONCLUSIONS

The crux of Turing's game is the machine's intellectual capacity to respond satisfactorily to unrestricted questions put by male or female interrogators. The authors oppose the idea that the machine in a Turing test should *imitate a man pretending to be a woman*, because it restricts the machine-human comparison test to a dependency on stereotypical female-male views on societal roles. Nonetheless gender concerns should be incorporated in the development of AI. More research is needed to find if embodied carers and companions or virtual assistants are accepted more as genderless, or with female or male, including as part of future healthcare.

REFERENCES

Artificial Solutions. *The Top Traits of Intelligent Virtual Assistants*. White Paper. Available here: http://www.artificial-solutions.com/about-artificial-solutions/resources/registered-whitepapers/ accessed 16.10.15

Clarey, C., 2009. 'Gender Test after a Gold-medal Finish'. http://www.nytimes.com/2009/08/20/sports/20runner.html accessed 11.9.15

CompanionAble, 2012. Integrated Cognitive Assistive and Domotic Companion- Robotic Assistants for Ability and Security. EU 7th Framework Programme (FP7): http://www.companionable.net/ accessed 15.9.15

Copeland, J. and Proudfoot, D., 2008. Turing's Test: A Philosophical and Historical Guide. In (Eds) R. Epstein, Roberts, G. and Beber, G. *Parsing the Turing Test: Philosophical and Methodological Issues in the Quest*

for the Thinking Computer. Springer, USA: pp 119 - 138.

Dattaro, L., 2015. Bot looks like a lady. Should robots have gender? *Slate*, available here: http://www.slate.com/articles/technology/future_tense/2015/02/robot_gender_is_it_bad_for_human_women.single.html accessed 15.12.15

Dennett, D. C., 2004. Can Machines Think? In (Ed) Shieber, S. *The Turing Test: Verbal Behavior as the Hallmark of Intelligence*. MIT Press: UK: pp 269-292

Dreger, A.D., 2010. Sex Typing for Sport. *Hastings Center for Report 40.* No. 2, March: pp 22-24

Gender Spectrum 2015. Understanding Gender. https://www.genderspectrum.org/quick-links/understanding-gender/ accessed 07.09.15

Genova, J., 1994. Turing's Sexual Guessing Game. *Social Epistemology*.Vol. 8, pp 313-326

Harnad, S., and Scherzer, P., 2008. First Scale up to the Robotic Turing test, then worry about feeling. *Artificial Intelligence in Medicine*. Vol. 44, Issue 2: pp 83-89

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. 972-977

Henry, B., 2014. Imaginaries of the Global Age. "Golem and others" in the post-human condition. *Politics & Society*, Vol. 2, 221-246

Hodges, A., 1992. *Alan Turing: the Enigma*. Vintage Books, London

IBT, 2015. Caitlyn Jenner Still Unsure About Sexual Preference. *International Business Times*. http://www.ibtimes.com/caitlyn-jenner-still-unsure-about-sexual-preference-1948066 2 June 2015

Lassègue, J., 1996. What kind of Turing Test did Turing have in mind? *Tekhnema 3/ A Touch of memory/Spring*. http://tekhnema.free.fr/3Lasseguearticle.htm accessed 11.9.15

Newman, L.K., 2002. Sex, Gender and Culture: Issues in the Definition, Assessment and Treatment of Gender Identity Disorder. *Clinical Child Psychology & Psychiatry*. Vol 7 (3), 352-359

Piccinini, G., 2000. Turing's Rules for the Imitation Game. In Moor, J.H. (Ed), *The Turing Test – the Elusive Standard of Artificial Intelligence* (2003) Kluwer, Dordrecht, The Netherlands, pp 111-119

Purtill, R.L., 1971. Beating the Imitation Game. *Mind*. Vol. 80, 290-294

Siegel, M., Breazeal, C., and Norton, M.I., 2009. Persuasive Robotics: The influence of robot gender on human behavior. *IEEE International conference on intelligent robots and systems* (IROS), 10-15 Oct, 2563-2568

Shah, H., 2014. Emotions of Alan Turing: The boy who explained Einstein's Theory of Relativity aged 15½ for his mother. *International Journal of Synthetic Emotions*, Vol 5(1), 23-30

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., and Warwick, K., forthcoming. Distinguishing AI from Male/Female Dialogue. *ICAART2016*, Rome, 24-26 February

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 *2nd Towards a Comprehensive Intelligence test (TCIT) symposium at AISB 2011*, University of York, 5 April

Shah, H., 2010. Deception detection and machine intelligence in practical Turing tests. PhD thesis, University of Reading, UK

Shah, H., and Warwick, K., 2010. Testing Turing's parallel-paired imitation game. *Kybernetes*, Vol. 39 (3), pp. 449-465

Sterrett, S., 2000. Sterrett, S. G. (2000). 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

Telegraph, 2015. Caster Semenya returns to running. http://www.telegraph.co.uk/news/worldnews/africaandindianocean/southafrica/11494602/Caster-Semenya-returns-to-running.html 25 March 2015

Turing, A.M., 1952 in Turing, A.M., Braithwaite, R., Jefferson, G. and Newman, M. Can Automatic Calculating Machines Be Said to Think? Transcript of 1952 BBC radio broadcast, in

Turing, A.M. 1950. Computing Machinery and Intelligence. *MIND*, Vol 59 (236), pp. 433-460

Universal, 2014. *Ex Machina* psychological thriller. http://www.universalpictures.co.uk/theatrical/ex-machina accessed 14.9.15

Vanity Fair, 2015. Call Me Caitlyn. July 2015 issue. http://www.vanityfair.com/hollywood/2015/06/caitlyn-jenner-photos-interview-buzz-bissinger accessed 10.9.2015

Warwick, K. and Shah, H., 2015. Human Misidentification in Turing Tests. *Journal of Experimental and Theoretical Artificial Intelligence*. Vol 27(2), 123-135

Warwick, K., and Shah, H., 2014c. Outwitted by the Hidden: Unsure Emotions. *International Journal of Synthetic Emotions*, Vol 5(1), 46-59

Warwick, K., and Shah, H., 2014b. Effects of Lying in Practical Turing tests. *AI & Society*, DOI: 10.1007/s00146-013-0534-3

Warwick, K., and Shah, H., 2014a. Good machine performance in Turing's imitation game. *IEEE Transactions on Computational Intelligence and AI in Games* 6(3), 289-299.

Westbrook, L. and Schilt, K., 2014 Doing Gender, Determining Gender. *Gender & Society*, Vol 28 (10), 32-57