

(Note to students: This sample policy brief should be used as a rough guide to give you a sense of the style, content and format. You should not use it as a rigid guide for what you are to hand in, as your assignment will likely differ slightly in terms of word length, content, and so on.)

Policy Brief - Granting Rights to Robots

Policy Topic: This government proposes to extend basic human rights, as defined in The Canadian Charter of Rights and Freedoms, to robots capable of passing the Turing Test (as originally described by Alan Turing).

I) Background and Description

The robot in question is a metal humanoid robot known as C-3PO. The robot C-3PO is equipped with proficient abilities in language and communication, and can adequately and appropriately respond to questions or concerns. In determining whether or not robots like C-3PO can think and should be considered persons, Turing suggests we test them with the imitation game, because “The question and answer method seems to be suitable for introducing almost any one of the fields of human endeavour that we wish to include” (Turing 4). Turing argues the question and answer method is effective because it covers many topics, spanning from poetry to mathematics, ensuring the broadest and most accurate testing. In order to be successful at the imitation game, robots must be able to answer questions in a way that makes them indistinguishable from a human being. It is recommended that the questioning be done over an instant messaging software program to ensure there are no biases present in the testing process, as “we do not wish to penalise the machine for its inability to shine in beauty competitions” (Turing 4). This method of testing ensures that it is C-3PO’s cognitive abilities that defines him as a person, therefore any entity with similar cognitive abilities will be considered a person regardless of whether or not they have a body.

II) Ethical Implications of the Policy Proposal

1) The Abolishment of Robot Servitude

Typically, a robot's purpose is to serve its human owner. However, once this legislation has passed, robots will have the freedom to discontinue any prior servitude, thus negating any purpose they once had. For example, Lin explains that "Robots are often tasked to perform the "three Ds", that is, jobs that are dull, dirty, or dangerous" (Lin 944), or in other words, tasks that humans are unwilling to do. Once granted rights and freedoms robots will no longer be required to perform these duties and will no longer hold any practical purpose. Without a practical purpose, there will be no economic advantage to creating robots capable of passing the Turing Test, which will endanger the future existence of the robots to whom this legislation is trying to protect. However, there is a solution to this ethical dilemma.

Robots can be designed in a way that will result in them performing practical duties willingly. The key to achieving this endeavor is to design robots in a way that they derive pleasure from performing their duties. For example, while the process of evolution has resulted in humans deriving pleasure from eating certain foods, robots could be designed in a similar way to derive pleasure from performing practical tasks. Peterson explains how engineers "could design APs to get their comparable reward rush instead from the look and smell of freshly-folded laundry" (Peterson 2). This would allow for society to benefit from robots continuing to perform their practical duties, such as folding laundry, without infringing upon the rights granted to them by this legislation.

2) Advancing the Morality of Society

Extending basic human rights to robots capable of passing the Turing Test will raise the overall level of morality and empathy within society. In order to pass the Turing Test, robots like

C-3PO must have advanced communication skills which will inevitably enable them to create emotional connections with human beings, putting them into the category of social robots. Instead of allowing these robots to be abused, Levy argues that “our own behaviour towards those robots [should be] an example of how one should treat other human beings” (Levy 214). Essentially, Levy is claiming that by protecting the rights of social robots, society will be setting a better example for how animals and humans should be treated. Similarly, Darling suggests that granting robots rights “may also prevent desensitization towards actual living creatures” (Darling 19), which will therefore create a more empathetic population. For example, if children are exposed to the abuse of social robots, this could desensitize them to the consequences of physical abuse, which could jeopardize the rights of animals and humans and lower the overall moral standing of society.

It should also be noted that Darling includes many types of low functioning robots that have very little resemblance to humans or animals in her argument, which could be a basis for criticism. However, the robots to whom this legislation is concerned with are higher functioning social robots. Although these robots may not feel pain, they may still have the self-awareness and intelligence to fear death/destruction. Thus, the repercussions of allowing social robots such as C-3PO to be abused and suffer, will lead to the degradation of society’s morals and values, which is why these robots should be granted rights and protection.

III) Normative Recommendations

1) A More Just Turing Test

One revision that should be made to the legislation is the implementation of a more just version of the Turing Test, than that which was originally proposed by Turing. The original test challenges the participant to fool the interpreter into thinking the participant is a human being of

average intelligence. However, many human beings with lower levels of intelligence such as children or cognitively handicapped people are still considered persons and are protected under The Canadian Charter of Rights and Freedoms despite their inability to pass the Turing Test. It is therefore unfairly discriminatory to revoke the rights of robots unable to pass the Turing Test, while still granting rights to humans who are unable to pass. Similarly, Leiber makes this argument in his book, *Can Animals and Machines Be Persons*, when Godwin argues that Washoe “behaves more intelligently and alertly than severely retarded humans. If the humans [. . .] are 100 percent persons, I think Washoe is 100 percent too” (Leiber 38-39). A fairer Turing Test would allow for robots of similar intelligence to Washoe to be able to pass. For instance, if the Turing Test were modified so that the participant had to impersonate a human child, rather than an adult, then robots less cognitively inclined than C-3PO such as the weaponized droids used by the empire, could still gain rights and freedoms. This change in legislation is important because robots unable to pass the traditional Turing Test may still be higher functioning social robots which according to Darling’s argument mentioned earlier, should still be protected for the sake of society’s moral standing.

2) Requiring a Physical Embodiment

In addition to having to pass the Turing Test, it is advisable that another clause be added to the legislation to require robots to have a physical entity/body in order to gain rights and freedoms. A physical embodiment is essential for a robot to interact with the world, have a greater sense of self awareness, and to gain and learn from physical experiences. Even AL from Leiber’s *Can Animals and Machine Be Persons* admits that his identity is not solely based on his central processing units but “also included my “peripherals”—my sensory and motor units, which themselves have acquired quite individual memories” (Leiber 57-58). These

characteristics differentiate robots such as AL or C-3PO from intelligent software programs in important ways.

For instance, defending the rights of physically embodied robots is both easier to legally enforce and is more beneficial to the morality of society, than the protecting a software program. Darling emphasizes this point, insisting that “A social robot is a physically embodied, autonomous agent” (Darling 2), suggesting that the abuse against software robots is not to the same caliber as abuse against embodied robots is. Thus, protecting the rights of software robots will not have the same moral benefits to society as protecting physical robots will. Furthermore, there is the instrumental challenge in defending the rights of software based robots, when their very existence could be eliminated by the delete button, or by a computer virus. It is therefore advisable, for both moral and instrumental reasons, to change the legislation to include only physically embodied robots.

IV) Possible Objection and Response

An objection could be made towards the use of the just Turing Test as described in section III-1 in determining whether a robot is intelligent or not. Searle makes this objection through comparing the Turing Test to the task of translating English to Mandarin. Searle argues that “there could be two "systems," both of which pass the Turing test, but only one of which understands” (Searle 6), and similarly, that there are two systems in translating English to Mandarin. The first system is a human who is fluent in both languages, and can translate the one to the other accordingly; this is compared to a human being passing the Turing Test. The second system is a human who is able to perform the translations using an intricate set of instruction manuals designed to allow him to translate the languages without having to learn or understand Mandarin; this system is compared to a computer passing the Turing Test. Thus, Searle argues

that the Turing Test is an inadequate method of testing intelligence because a computer could theoretically pass the test without having any genuine cognitive understanding of what they are saying.

The issue with this objection is that it could be used on both humans and computers alike. For instance Godwin, in Leiber's *Can Animals and Machines Be Persons*, uses the objection to ask Goodman, "How am I to know that you don't have inside you a tiny Chinese man who understands the Chinese rules he follows in converting them [. . .] into English output symbols?" (Leiber 51). Goodman responds by explaining all of his feelings and experiences, after which AL responds by similarly explaining all of his feelings and experiences. Furthermore, AL argues that the thoughts produced by his many processing chips are no less genuine than thoughts being produced by many neurons. Thus, because there is no way to distinguish genuine cognitive understanding from artificial understanding in either humans or robots, it is only fair to judge an entity's intelligence by the understanding they convey through language; which is why the Turing Test is a valid test for intelligence.

Works Cited

- Darling, Kate. "Extending Legal Rights to Social Robots." *SSRN Electronic Journal SSRN Journal* (2014): 1-23. Web. 12 Oct. 2015.
- Leiber, Justin. *Can Animals and Machines Be Persons? A Dialogue*. Indianapolis: Hackett Pub., 1985. 38-59. Print.
- Levy, David. "The Ethical Treatment of Artificially Conscious Robots." *International Journal of Social Robotics* (2009): 209-16. Web. 14 Oct. 2015.
- Lin, Patrick, Keith Abney, and George Bekey. "Robot Ethics: Mapping the Issues for Mechanized World." *Artificial Intelligence* (2011): 942-49. Elsevier. Web. 16 Oct. 2015.
- Peterson, Stephen. "Designing People to Serve." (2010). Niagara University. Web. 17 Oct. 2015.
- Searle, John. R. (1980) Minds, brains, and programs. *Behavioral and Brain Sciences* 3 (3): 417-457. Web. 15 Oct. 2015.
- Turing, A.M. "Computing Machinery and Intelligence." *Mind a Quarterly Review of Psychology and Philosophy*. Oxford University Press, 1 Oct. 1950. Web. 25 Oct. 2015.