

FRANCISCO, Augusto – What if Alan Turing, Masahiro Mori, and John Searle employed a humanoid robot in a *Gedankenexperiment*?. *Configurações: Revista de Ciências Sociais* [Em linha]. 37 (2026) 205-228. ISSN 2182-7419. <https://doi.org/10.21814/config.7016>

VÁRIA | JUNHO 2026

WHAT IF ALAN TURING, MASAHIRO MORI, AND JOHN SEARLE EMPLOYED A HUMANOID ROBOT IN A GEDANKENEXPERIMENT?*

AUGUSTO FRANCISCOⁱ

Technical University of Dresden

ABSTRACT

Critical constructionist theory employs three interconnected dimensions (real, symbolic and imaginary) to clarify the emergence of social constructions. The use of humanoid robots in scientific experiments is affected by these dimensions. By considering this perspective, we explore the arguments of Turing, Searle, and Mori to develop a new thought experiment. The humanoid robot acts hereby as a platform for the interaction of the imaginary, the symbolic and the real. The functioning of the technological context strongly intertwines with its “interior” and “exterior”. Consequently, it is not scientifically valid to label individual technologies (without context) as superior or inferior.

KEYWORDS: Turing Test; Chinese Room; Uncanny Valley; humanoid robotics; technology

ⁱ acfrancisco@daad-alumni.de | ORCID: 0000-0003-1602-8326.

* I could not have come up with any thoughts about this fascinating technological tool, the humanoid robot, without the participation of roboticists in my research. I am grateful to both them and their institutions for granting me the privilege of being with them. The subjects in my research are shielded by anonymity. I would like to extend my thanks to friends and colleagues who offered comments and criticisms. Specifically for Marie Charbonneau and Nuno Guedelha. I hope I have accurately incorporated their suggestions, thoughts, and corrections into this work. I would like to extend a special thank you to Lynn and Hanna Nohe.

RESUMO**E SE ALAN TURING, MASAHIRO MORI E JOHN SEARLE EMPREGASSEM UM ROBÔ HUMANÓIDE EM UM *GEDANKENEXPERIMENT*?**

A teoria construcionista crítica emprega três dimensões interligadas (real, simbólico e imaginário) para esclarecer a emergência das construções sociais. O uso de robôs humanoides em experimentos científicos é influenciado por essas dimensões. Ao considerar esta perspectiva, exploramos os argumentos de Turing, Searle e Mori para desenvolver um novo experimento mental. O robô humanoide atua assim como uma plataforma para a interação do imaginário, do simbólico e do real. O funcionamento do contexto tecnológico entrelaça-se fortemente com o seu “interior” e “exterior”. Consequentemente, não é cientificamente válido rotular tecnologias específicas (sem contexto) como superiores ou inferiores.

KEYWORDS: Teste de Turing; Quarto Chinês; Vale da Estranheza; robótica humanóide; tecnologia

RESUMÉ**ET SI ALAN TURING, MASAHIRO MORI ET JOHN SEARLE UTILISAIENT UN ROBOT HUMANOÏDE DANS UNE *GEDANKENEXPERIMENT* ?**

La théorie constructionniste critique emploie trois dimensions interconnectées (réel, symbolique et imaginaire) pour élucider l'émergence des constructions sociales. L'utilisation de robots humanoïdes dans des expériences scientifiques est influencée par ces dimensions. En considérant cette perspective, nous explorons les arguments de Turing, Searle et Mori pour développer une nouvelle expérience de pensée. Le robot humanoïde agit ainsi comme une plateforme d'interaction entre l'imaginaire, le symbolique et le réel. Le fonctionnement du contexte technologique est lié à son « intérieur » et à son « extérieur ». Il n'est pas scientifiquement valide de qualifier des technologies spécifiques (sans contexte) de supérieures ou d'inférieures.

MOTS-CLÉS: Test de Turing ; Chambre Chinoise ; Vallée de l'Étrange ; robotique humanoïde ; technologie

Introduction

In this paper, I will show sociologically that the functioning of humanoid robots depends on three essential interconnected dimensions: imaginary, symbolic and real. The *symbolic* code and design operate a *real* (robotic) “body”, which is *imaginarily* believed to be (like) a human. In the first part, I will draw a comparison between (1) Alan Turing’s “Turing Test” and the *imaginary*, (2) John Searle’s “Chinese Room” and the *symbolic*, (3) the experimental technology “humanoid robot” and the *real*, and lastly (4) Masahiro Mori’s “Uncanny Valley” and the relationship between imaginary, symbolic, and real. I will adopt the critical constructionist viewpoint that humanoid robot technology is a social tool for collecting data within the scientific system. In the second part, a new *Gedankenexperiment*¹ is described wherein fictitious Turing, Mori, and Searle examine their arguments through a collective effort with the assistance of a humanoid robot (called *h_Mori*). In the third part, I will explore the intersections of the three dimensions (imaginary, symbolic, and real), the three concrete arguments (Turing, Mori, and Searle), and the three fictitious arguments (*Gedankenexperiment*) with *h_Mori*. Finally, I will make some concluding remarks.

1. INVESTIGATING HUMANOID ROBOTS AS A MEDIUM FOR TURING, MORI, AND SEARLE’S EXPERIMENTS

Following my ethnography in robotic laboratories (not yet published), I have been pondering the imaginative possibility of Alan Turing, Masahiro Mori, and John Searle employing a humanoid robot to mediate their arguments in experimental practice. I am intrigued by how the tests/experiments/hypotheses of these three prominent characters and the humanoid character play out the imaginary, symbolic, and real scene. I intend to link this scene to the concept of critical constructionism by Alípio De Sousa Filho (2019) and technology by Jost Halfmann (1996).

Critical constructionism plays a crucial role in unifying my assumptions on humanoid technology. As sociologists, we need to move past

¹ “Gedankenexperiment” is a German word meaning “thought experiment”. It is employed by Searle himself (1980, p. 417) to explain his model for analyzing artificial intelligence.

endless theoretical debates that pit theories against each other without any practical benefits. Criticizing the “narcissism of minor theoretical differences”² can positively impact sociology. So, it is reasonable to consider critical constructionist thinking (De Sousa Filho, 2019, pp. 13-14):

As an underlying theory of the human sciences, I view critical constructionist theory, and its consequent deconstructionism, as the result of work of researchers and authors who differ among themselves - including some who did not even claim the title ‘constructionist’ or ‘deconstructionist’. Instead, critical constructionist theory serves to synthesize thoughts, discoveries and conclusions in research and studies in the fields of anthropology, sociology, history, philosophical conceptions, linguistic theory, psychology, and psychoanalysis. Underlying major perspectives in the human sciences - such as functionalism, structuralism, Marxism, rational utilitarianism, symbolic interactionism, phenomenology, critical realism, critical theory, pragmatism, dispositionalism, figurationist theory and complexity - we find elements of *critical constructionism* in these analyses approaches and perspectives.

De Sousa Filho’s (2019) concept of critical constructionism is based on theories that explain how the *imaginary* arises from the initial image we have of our neighbors and their actions, and how this image underlies social life³. In turn, the *symbolic* encompasses language that assigns meaning to the world and allows individuals to discuss their social position⁴. Finally, the *real*, which is impossible to grasp, marks the border of everything that

2 Freud defined *narcissism of minor differences* which we can apply to sociological “Schools”: “(...) gerade benachbarte und einander auch sonst nahe stehende Gemeinschaften sich gegenseitig beföhden und verspotten” (1999b, p. 473). Or in the English version: “(...) it is precisely communities with adjoining territories, and related to each other in other ways as well, who are engaged in constant feuds and in ridiculing each other” (1961, p. 114).

3 De Sousa Filho builds on the influential ideas of Castoriadis (1975). We attempt to incorporate this understanding into humanoid robotics analysis. From the perspective of the imaginary, we have included an additional critical constructionist author in our thesis with the support of DeSousa Filho: Lacan (1978).

4 The symbolic is essential to his thinking, just like the imaginary. Significant authors who work with the symbolic: Durkheim (1990), Lévi-Strauss (2017), Lacan (1975, 1978), Castoriadis (1975), and Bourdieu (1998).

exists and, simultaneously, is the societal “ballast” (p. 41)⁵ of every social order. *Imaginary*, in terms of sociology of technology, could hypothetically refer to the belief in the image of technology’s functioning. When it comes to humanoid robots, it would refer to believing in the functioning of a humanlike image. The *symbolic* could represent the set of operations that give names to the technological process. In the robotic scenario, it would be what “animates” the humanlike design and programming code of its structure. The *real* could be the formless mass of natural events within the technological object. In the current humanoid scenario, the amorphous (robotic) carcass would represent the real and serve social functions in laboratory settings from an imaginary and symbolic perspective.

In turn, Halfmann’s (1996) concept of technology is influenced by Niklas Luhmann’s (1984) systems theory. Halfmann’s (1996) proposal states that technology serves as a *medium of communication* for social systems and a medium of consciousness for psychic systems. What happens if the technology does not work, because it is either broken or not ready to be used? It is not a *medium* in these cases, but rather an *installation*. Thus, the installation can be fixed and transformed into a medium; or transformed into a medium if it has never been a technology. From this perspective, technologies are constantly connected and flowing between installation and medium. In order to operate, a humanoid robot must possess multiple technologies and be able to “interact” with diverse other technologies (flat floor, doors, obstacles...). For instance, they are equipped with sensors, wires and metals, and have the ability to operate other technologies like door handles. Halfmann’s sociology of technology could also be compared to the dimensions of the imaginary (our expectation is that the installation will become a medium *or not*) (Halfmann, 1996, p. 87, 120), the symbolic (combination of cause-and-effect operations) (p. 99, pp. 119–21), and the real (the *nicht-sinnhafte* nature of technology) (p. 88, 116)⁶.

5 The concept of the “real” plays a central role in his work. His thinking about the real has been influenced by important authors such as: Lacan (1978, 1981), and Castoriadis (1975).

6 The approximate translation of the German word *nicht-sinnhaft* is “non-meaningful”. *Sinn* is the German word for “meaning”. *Sinnhaft* means “meaningful”. According to the author (p. 116): “Sachliche Systeme wie Technik bestehen aus physikalischen oder chemischen Elementen, die grundsätzlich nicht-sinnhaft sind” (translated by us: “Material systems such as technology consist of physical or chemical elements that are fundamentally non-meaningful”).

In the case of humanoid robots, the expectation is hypothetically that the human likeness of the robot would work through well-debugged operations of codes that work on raw matter which is, at heart, *nicht-sinnhaft*. Our premise is that technologies (humanoid robotics) are always in “danger” of regressing to the installation state. However, humanoid robots are a relatively successful scientific medium for gathering data within a laboratory environment. The technology is in constant tension as the robot’s data is almost always confused with its “humanlike” characteristic.

After provisionally defining the analytical concepts of this work, it is crucial to review each original test/experiment/hypothesis, at least in its basic form, in order to then consider the topic from the perspective of humanoid robots, as a form of technology.

The original Turing test attempted to determine whether a computer can exhibit humanlike thinking (Turing, 1950). Turing led us to believe that one’s uncertainty about whether or not a machine could replace a person in the exchanging of written messages, was supported by the machine’s ability to imitate thought. In order to avoid any interference from the experimental set, messages would have to be typed and transmitted without contact during the test (e.g., voice-free contact is necessary to maintain the test’s credibility, given that the computer’s voice would reveal itself). Turing is in 1950! Then, an interrogator (C) should be unable to distinguish if a person position A is temporarily occupied by a computer. In addition, C also gets written messages from person B, who is actually a real person. An imitation game existed prior to Turing: C had to determine if A and B were male or female. In the original Turing test, if C is unable to differentiate A as a machine, then the machine can imitate the person position A. Turing believes that the machine in position A has the ability to mimic thought. In order to consolidate the imitation test, Turing presents a series of counterarguments to the criticisms against his thinking machine model.

Searle’s original proposition of the “Chinese room” argues that a computer can pass the Turing test despite lacking understanding of message meaning (Searle, 1980). Understanding the message’s content is necessary for thinking. According to him, the strong artificial intelligence

thesis cannot be achieved solely through computational symbolic repetition. Searle presents the following *Gedankenexperiment* to support his argument. In a closed room with no contact to the outside world, Searle performs symbolic operations similar to those on a computer, following an instruction batch in English. Inside this room, he is given Chinese information from outside, not even recognizing it as a language. The unknown information needs to be manipulated using those instructions symbolically, formally, so that certain symbols are connected to others and sent out of the room. Instantiation is the term he uses for this process (p. 418). The Chinese speaker outside the room cannot tell that Searle does not speak Chinese because of his coherent responses. This demonstrates that he lacks understanding of the meaning of Chinese symbols when manipulating formal rules. Similarly, the computer lacks understanding of the meaning of these symbols. As a result, neither of them can think. Furthermore, Searle states that if he receives information in English and responds in English, both he and the other person will understand the content, but the computer will not. Searle further explains a sequence of counterarguments against those who criticize him.

Lastly, Mori did not directly create any test or *Gedankenexperiment*, but he did speculate on the concept of “Human Likeness” in objects, including (humanoid) robots. *His original thesis centers on the idea of the “uncanny valley”* (Mori, 2012). He suggests that the more objects resemble humans, the stronger our affinity towards them. However, this affinity has a limit. Beyond this point, increased human likeness results in decreased affinity and we start to feel an uncanny effect. When the object’s variable “movement” is a factor, decreased affinity is intensified by the movement of the humanlike object. However, when the object is given much more “humanlike” attributes again, our affinity increases once more. In his graph, Mori uses the x-axis to represent Human Likeness and the y-axis to represent affinity. He calls this decrease in affinity after a rise in human likeness the “uncanny valley”, represented by a V-shaped graph (resembling a valley). Affinity drops dramatically, reaching the valley floor, and then increases as the object is given more “humanlike” attributes again. To avoid the uncanny valley, Mori recommended designing robots that do not closely resemble humans.

The purpose of this paper is neither to investigate artificial intelligence nor to discredit the mentioned tests. It is the opposite: they can all be validated. My hypothesis: they can only be validated if done in conjunction with a common technology that works. For successful technology mediation, the imaginary, the symbolic, and the real must be “intricately” connected, allowing seamless communication. I suggest a joint critical constructionist approach in conducting the experiments mentioned above, incorporating humanoid robotic technology. In the next chapter, I will introduce the *WG Communication and Consciousness*, a new *Gedankenexperiment* where I will redo them (Turing, Mori and Searle) together.

If we were to place the Turing Test, the Chinese Room, the humanoid robot and the Uncanny Valley in the hypothetical scenario I am suggesting, it would resemble (1) an *imaginary* test where we cannot differentiate between a humanoid robot and a human; (2) a *symbolic* experiment where we (including the robot) are unaware of the instructions we execute; (3) a *real Nicht-Sinnhaftes*⁷ (real non-meaningful thing) where we are confronted with an inherent lack of knowledge about what occurs within the robot; and (4) an *interaction* between the imaginary, the symbolic and the real where the humanoid robot’s humanlike characteristic is successful and its code is symbolically well structured around the real, resulting in affinity.

2. WG COMMUNICATION AND CONSCIOUSNESS

WG is the abbreviated form of the German term *Wohngemeinschaft*, which translates to “shared flat”. In our experiment the best friends Turing, Searle, and Mori live together in the *WG Communication and Consciousness*. There is a room exclusively for Turing, who represents the imaginary, while Mori and Searle share another room, representing the symbolic. Their interests lie in conducting real and thought experiments. The friends are supportive and communicative, highly intelligent, and committed to helping each other. They will collectively test their arguments today and need a humanoid robot for this purpose, which

⁷ The adjective *nicht-sinnhaft* transformed into a noun.

represents the real. Mori will build it to resemble himself and call it *humanoid Mori* or *h_Mori*. It will occupy the third and last room of the *WG*.

The rooms are organized in a manner that resembles a pizza sliced in three. At the center, there is a communal area with a round table. Spaces are assigned names. *Intersection* is the name of the central area. Turing is in a room called *Imaginary*. The shared room of Mori and Searle is called *Symbolic*. The room *h_Mori* will stay in is named *Real*⁸. The *Real* was set up with technological requirements for *h_Mori* by our three intellectuals⁹. Participants must avoid entering other people's rooms or gathering at the *Intersection* during the experiment to satisfy the Turing criterion¹⁰. Thus, Turing is not allowed to enter Mori and Searle's room or *h_Mori*'s room. Mori and Searle are only permitted to enter their own room. The same rules also apply to *h_Mori*.

Helped by Turing and Searle, Mori builds *h_Mori* for experiments in the *Intersection* (before starting the experiments). Mori says:

— The graph's X-axis displays Human Likeness, ranging from 0% to 100%, with 0% having no relation to humans and 100% representing a healthy human¹¹. A humanoid robot, for instance, would reach the range of 30% to 50%.

And, in fact, with the help of his friends, Mori manages to reach 50% Human Likeness. Hence, *h_Mori* shares many similarities with Mori, including the ability to enter and exit the *Real* using his own legs and hands. Additionally, it contains a translation program for English to Japanese¹² and vice versa, along with an intelligent chat program to engage in written conversations with Turing. Consequently, among other tasks,

8 Capitalized "Imaginary", "Symbolic", and "Real" represent occupied rooms, while lower-case "imaginary", "symbolic", and "real" represent critical constructionist concepts. While they can be confused with each other, the rooms embody ideal delimitations of these concepts. Weber's (1904) notion of the ideal type is not very different from this.

9 However, the *Real* had previous issues that were temporarily fixed.

10 "(...) condition which prevents the interrogator from seeing or touching the other competitors, or hearing their voices." (Turing, 1950, p. 434).

11 In a future paper, I will criticize this ideological graph/idea.

12 Mori is Japanese.

it can read letters as well as open and seal envelopes. With everyone satisfied with the outcome, they begin preparing the experiments.

HOW WILL THE JOINT EXPERIMENTS BE CARRIED OUT?

Turing's experiment in the Imaginary requires writing two identical notes (with similar handwriting) sealed in two envelopes to initiate the conversation. After leaving the envelopes on the Intersection table, he will go back to his room. The experiment's sequence will be randomized. This means Turing will write the notes while Mori flips a coin to decide if *h_Mori* will be involved in the experiment. The conditions of the Turing test should reflect uncertainty about the authorship of these notes: Mori or *h_Mori*? It is similar to a blind test. Either only Mori will respond to the two notes or both Mori and *h_Mori* will¹³.

The organization of Mori and Searle's room enables the continuation of the (imaginary) "Turing test" in the (symbolic) "Chinese Room"¹⁴ experiment and the articulation of *h_Mori*'s technological (real) carcass. Thus, the Symbolic will be split into two sections with "no" communication, as the door connecting them will stay locked. Nevertheless, there is a tiny opening in this division of the Symbolic where an envelope can be transferred from one section to another. Searle will occupy the section facing the door of the Intersection. Mori will be in the section of the room that does not connect with the Intersection. The subsequent events are contingent upon the outcome of the blind "coin flipping" test.

*MORI AND H_MORI WILL BOTH RESPOND TO THE TWO NOTES*¹⁵

The experiment was designed to ensure that everyone knew when it was time to exit their rooms without meeting: inside, above the door in each room,

¹³ The "Chinese Room" test would become impossible if *h_Mori* responded to both notes.

¹⁴ Our Searle thought experiment will be referred to as the "Japanese Room" since Mori is Japanese and speaks Japanese.

¹⁵ My colleague Marie Charbonneau warned me about the possibility of modern generative AI technology producing multiple outputs for a single input. To make my model simpler, I had to set the condition that the two notes are highly alike. In simpler terms, both the Japanese Room and *h_Mori* would produce only one possible output for each input (Turing's written note).

there's an indicator light that changes from red to green at this moment. Turing gets the green light and soon others will get the green light too. After leaving the two envelopes on the table in the Intersection, Turing will return to the Imaginary, and then Searle will get the green light, take the envelopes and go back to the Symbolic. The envelopes will be passed from Searle to Mori. Mori will proceed in the following manner; he will set aside a sealed envelope for *h_Mori*, open the second envelope, extract the note, and use the translation program also used by *h_Mori* to translate it into Japanese on his laptop.

Next, the Japanese translation will be written by hand on a fresh note and given to Searle. In turn, Searle will apply the same intelligent chat instructions as *h_Mori*. These instructions are provided in the instruction book for processing Japanese symbols (there is no chance that Searle will understand any of the scribbles on the note). Following that, he will write down the result on a new note and give it to Mori, who will use the aforementioned *h_Mori* program to translate it into English. This is the response to Turing's note, but Searle will not know what it means. Mori will write this reply on a new note, seal it in an envelope, and then give both envelopes to Searle, who will leave them on the Intersection table. Upon Searle's return to the Symbolic, *h_Mori* will get the green light and leave the Real in order to collect Turing's envelope¹⁶. After picking up the envelope with Turing's note, it will bring it to the Real and respond, as he has been programmed to do. It will then leave it on the table in the Intersection.

Turing will take the two response envelopes and, if unsure of their authorship by *h_Mori* and Mori, he will reply to them, repeating the process for ten minutes of chat (ten minutes only account for reading and writing the messages between Turing, Mori, and *h_Mori*, excluding the time of letter deposition and processing by Searle)¹⁷. Unless he is already aware of *h_Mori*'s presence, *h_Mori* will successfully complete the Turing test. Evidently, in the event Turing discovers it is *h_Mori*, *h_Mori* will be unsuccessful. In turn, if the Turing test is successful, then the Japanese room is also successful. Searle will prove *h_Mori*'s lack of understanding. In a

16 The sensors of *h_Mori* can differentiate Turing's note by detecting the relatively heavier paper he uses.

17 According to Turing, a valid criterion could potentially be five minutes (1950, p. 442): "(...) an average interrogator will not have more than 70 per cent. chance of making the right identification after five minutes of questioning". In order to be more conservative in the methodological criteria, we will double this value.

hypothetical scenario where they are successful, it can be argued that h_Mori has successfully passed the “uncanny valley” in terms of writing.

ONLY MORI WILL RESPOND TO THE TWO NOTES

In this scenario, Searle will pick up the envelopes on the table and everything will play out practically the same as before. The only variation is that Mori will create two, not one, similar notes, once processed by Searle. To put it differently, h_Mori will not be involved in the experiment. Similarly to the previous case, Searle will place the two envelopes on the Intersection table. Turing will take the two envelopes and if h_Mori is not identified, the “game” will continue for ten minutes before Mori discloses that only humans are involved in the test. As a result, the Turing test will be “inconclusive” due to the absence of h_Mori. The “Japanese Room” will be “inconclusive” for the same reason. The Uncanny Valley will experience the same outcome. If Turing identifies h_Mori, it is considered a false negative.

FINAL PREPARATIONS FOR THE EXPERIMENT

Everyone goes to their rooms, including h_Mori. Mori asks himself:

— Will h_Mori fall into the uncanny valley with 50% of human likeness?¹⁸

Turing asks himself:

— What will happen if h_Mori takes the part of Mori in this game?¹⁹

Searle asks himself as well:

— Can I be simply an instantiation of h_Mori?²⁰

18 In the original text (Mori, 2012, p. 100): “We hope to design and build robots and prosthetic hands that will not fall into the uncanny valley”.

19 In the original text (Turing, 1950, p. 434): “What will happen when a machine takes the part of A in this game?”

20 In the original text (Searle, 1980, p. 418): “(...) I am simply an instantiation of the computer program.”

LET US START THE EXPERIMENT!

At the table in the Imaginary, Turing writes the following message on two notes: “Fire an interesting question at me!”. Next, he puts them into two envelopes and seals them. Mori’s coin- flipping results in heads, meaning *h_Mori* participates without Turing’s and Searle’s awareness. Green light. He walks out of the Imaginary, leaving the two envelopes on the table in the Intersection. He returns to his room.

After entering the Imaginary, Turing is followed by Searle who then notices the green light and exits the Symbolic, takes two envelopes, and returns to his room. When Searle enters the room, he passes the envelopes to Mori, who translates them. Mori gives Searle an envelope containing Japanese input, which is the Japanese translation of Turing’s note: “Fire an interesting question at me!”. Using this input data, Searle executes the formal intelligent chat rules (the exact same chat program as *h_Mori*). Searle has come to a result. He writes down the result and passes it to Mori, who translates it into English (the exact same translation program as *h_Mori*) and seals it in an envelope. Then, he gives the two sealed envelopes back to Searle. With the envelopes in hand, Searle then leaves the room and places them on the Intersection table. It can be inferred that Mori read and understood the content of the open note. However, Searle came across a sequence of indecipherable signs and consequently failed to grasp any meaning from them.

Once Searle returns to his room, *h_Mori* (with its sensors) is able to identify the green light, leave the Real and proceed to the table. Using its sensors, it selects the envelope with the heaviest weight and goes back to its room. Two minutes later, it exits the Real and places a sealed envelope on the table, before entering its room again. It remains uncertain how *h_Mori* proceeded in Real, including whether there is a note inside the envelope and, if so, what it says. It is presumed that *h_Mori* has managed the translation program from English to Japanese (and vice versa) and the intelligent chat program, both created by Mori.

After *h_Mori* goes into its room, Turing exits the Imaginary, collects the two envelopes and goes back to the Imaginary. He opens them and finds that the two notes have nearly identical handwriting to Mori’s: “Am I a

human?”. Turing is startled. As the two notes are virtually identical, he is unable to distinguish which one was written by a human and which one was potentially written by the humanoid. For this reason, he decides to respond to the two notes and continue the conversation: “You look like a human! Can you please tell me your name?”. There will be another round of the process. Green light. Turing to Searle, Searle to Mori, Mori to Searle, Searle to Mori again, Mori to Searle again, Searle to h_Mori and h_Mori to Turing. Once again, the notes have a similar style of writing: “Hey, it’s me, Mori, your friend. We’re friends, aren’t we?”.

It is Turing’s turn: “We are definitely friends! But we have different ways of thinking, don’t we?”

Mori/Searle and h_Mori: “We have different ways of thinking and we are friends.”

Turing: “If h-Mori and I are exchanging notes now, can I consider him to have passed the Turing test?”

Mori/Searle and h_Mori: “Do you think h_Mori is playing the game?”

Turing: “Actually, I’m completely unaware of that. Was the ‘Japanese Room’ successful?”

Mori/Searle and h_Mori: “That’s an excellent question. Searle is there, cut off from everyone else.”

Turing: “Let’s keep exchanging notes because I’m still uncertain if h_Mori is playing this game!”

Mori/Searle and h_Mori: “Let’s keep going, as humans do. Fire an interesting question at me!”

The conversation goes on for ten more minutes. Once all the experiments are empirically and officially proven by the three friends (all

data was recorded!), Turing decides to write in the next note to end them and take a break with a beer. Something unexpected unfolds. The indicator light that alternates between red and green no longer works. Despite that, Turing leaves the room to get the envelopes, but there is only one on the table. It is odd, so he decides to go back into the Imaginary and wait a while before coming out again. Once again, he goes out and nothing happens. Turing's hope remains. Those who are in the Imaginary always hold strong beliefs! He steps inside the room and waits a little longer. In turn, Searle needs to go out and collect the two envelopes. But the indicator light in Searle's room also no longer works. Thus, he leaves the room without knowing anything. Turing and Searle coincidentally depart from the Imaginary and the Symbolic simultaneously. At first, they gaze at each other in astonishment. After their conversation, they "get *real*" about issues concerning the indicator light and h_Mori. They contact Mori and collectively opt to enter the Real. As the Real is locked, breaking down the door becomes necessary.

Breaking down the door reveals this situation: h_Mori is repeatedly walking in a circular path in the Real, while making peculiar hand movements as if writing sideways in the air. For those not in the field of robotics, h_Mori likely embodies all the uncanny feelings. When h_Mori falls to the ground and spins its head like in the 1973 film "The Exorcist" (Friedkin, 1973), this feeling is further reinforced. However, our three intellectuals perceive it in a different way. The feeling a roboticist experiences when confronted with a malfunctioning robot is usually expressed as irritation. It is necessary to press the big red push button²¹ for an immediate shutdown of h_Mori.

At that moment, the friends begin their observation of the Real. Taking a closer look, they come to the realization that there are several issues in Real: the floor is wet, there's a leak from a cracked pipe, the dehumidification device is inefficient, h_Mori is holding a pen upside down, the table is cracked, the chair is shaking, and there's a broken pin in h_Mori's battery connection. On the table, there is a note from Turing with a smudge caused by a drip. The wetness of the floor caused

21 The red button is frequently used in robotics. Pressing it causes the robot to lose power and subsequently turn off.

h_Mori to step wrongly, resulting in a crooked foot. The battery is wet and in critical condition. Short circuits are a potential risk. Mori reviews h_Mori's software and identifies issues with the inverted pendulum and foot sensors, leading to erratic values. To sum up, there are many technical issues.

The repair will take a long time, so the experiments cannot be repeated today. In fact, they are already late for tonight's classes. Hannah Arendt will be teaching the first class on the banality of evil. Their goal is to apply Arendt's ideas to advocate for a manifesto condemning the banality of utilizing robots for sexual purposes. Kathleen Richardson is their source of inspiration²². Then, they will attend a class with Donna Haraway to acquire the skills necessary to compose a controversial manifesto. After the classes, they agree to go to the *Kneipe* (bar) to discuss the results of the experiment, while enjoying a beer from *Bolten Brauerei* called *Alt*. At the *Kneipe*, they find common ground on one point. The consistency of the "Turing test", Mori's "uncanny valley", and Searle's "Japanese room" are not the issue. In the *WG* scenario, all tests have proven successful, albeit temporarily. In fact, they conclude that the positive results are linked to the tests being intertwined. All rooms are important. The results are limited to when h_Mori is working. When it stops working, the question of a positive or negative outcome disappears. Neither a positive nor a negative outcome exists as a whole anymore. What remains are fragments of reality, in the form of data collected from various sources (sensors, actuators, cameras, etc)²³.

3. COMMUNITY LIFE: WG CAN ONLY EXIST THROUGH COEXISTENCE

In Germany, *WG* is a place where coexistence is emphasized. In a *WG*, you leave your room to socialize in common areas: you talk, share experiences, conduct experiments, and maybe even you all go to a bar to drink beer and discuss politics. This is why *WG Communication and Consciousness* was able to work so well with the coexistence of Turing, Searle, and Mori.

22 Available in: <https://campaignagainstsexrobots.org/>

23 The choices made by our intellectuals when it comes to defending and teaching women's rights, philosophical preferences, or indulging in sweets or beer are purely imaginative and exist solely within the literary context of this thought experiment. These alternatives are fictional. I hold immense respect for each and every one of them.

Turing's argument, centered on the "imaginable computer" (Turing, 1950, p. 436, 442), mirrors what scientists, philosophers, theologians, and ordinary people like us do: reside in the dimension of the imaginary. Based on my belief that I am interacting with a thinking entity, I assign a computer the role of a person's position "A" in the imitation game. "A" truly thinks - but from the dimension of the imaginary. It is a fact of the imaginary! If my conversation with "A" becomes convincingly coherent, it fills the void of uncertainty about him. The words spoken to me evoke an image of someone who, like us, thinks and uses language to find the right words for each occasion. Finding the right way to talk may involve mistakes and inconsistencies during the conversation itself, as Turing discussed in the mistake simulation (pp. 448-49).

When we consider a humanoid robot instead of a computer for "A", the imaginary becomes more vivid. This is because the humanlike image is the first one that makes sense in social life. Lacan coined the term "mirror stage" for this initial image in the formation of the "I" (ego) (Lacan, 1949). My own I is always rooted in the humanlike image of another I. Despite this, it continues to be an empty image. If this image only brings a sense of emptiness (*inanité*) (p. 449), it can be easily replaced by anything resembling my "I" - anything that looks like me, any humanlike image. Lacan established a connection between this image and the image of statues and automatons (p. 451) (humanoid robots were likely not considered because they were not a social reality in 1949). The image of others does not have to be "human" or made of flesh and blood. In fact, within itself, it can never be human since it is just an image. It is just a humanlike image. However, it consistently presents itself as a human, making it feel like *I am face-to-face with an I*. Therefore, from an imaginary standpoint, any humanlike image would be indistinguishable from a human. The first effect of the imaginary is established.

In the *WG Communication and Consciousness*, it is not uncommon for Turing to perceive h_Mori as a humanlike image during the belief of exchanging notes. The fact that Turing cannot distinguish between Mori and h_Mori implies that h_Mori *is* actually a humanlike image. Therefore, h_Mori meets the criteria of the Turing test for our WG. Initially, nothing appears to be wrong with that. The tendency to attribute

anthropomorphic qualities to humanlike images is distinctly human. As an example, it is said that Goethe was sensually influenced by a humanlike statue (Hersey, 2009, p. 6). Additionally, Freud (1999a) presents the case of Gradiva. Hoffmann's "Sandmann" (2004) is another example. Plato's Cave (2000, pp. 220-251) is the philosophical origin of this problem of humanlike images, where the cave wall displays such representations. The transition from an image to a humanlike image requires a symbolic operation.

At this point, the symbolic becomes relevant. The concept of the "Chinese Room", as described by Searle, is meaningful only within a framework that incorporates symbolic elements: "They enable me to correlate one set of formal symbols with another set of formal symbols" (Searle, 1980, p. 418). In the position of the computer that passed the "Turing test", Searle states: "I produce the answers by manipulating uninterpreted formal symbols" (p. 418).

Searle's imaginary identifies a thinking being in that humanoid image, but it is merely a collection of symbolic operations which hold no meaning for him. Although they are symbolic operations, a concatenation of symbols, if they do not make sense to me, they cannot represent thought as they do in normal situations. If they lack meaning for me and, as a result, I do not think of using them, then they also lack meaning for a computer, which consequently does not think. Consider the way we perform tasks automatically while our thoughts are focused elsewhere. I perform the symbolic operation of washing dishes while imagining a fictitious encounter between Turing, Mori, and Searle while enjoying an *Apfelstrudel*. In a way, it is not completely different from Lacan's thesis on the *chaîne signifiante* (signifying chain) (Lacan, 1966, pp. 11-41). Despite Lacan discussing the unconscious, his argument emphasizes the operation of symbols on me without my conscious understanding.

The symbols in the "Japanese Room" now remain linked to the signifiers, despite Searle's lack of knowledge about their meaning. The symbolic has power over him, a human, without his conscious understanding - this is the idea! Hence, the symbolic operations that used to be comprehensible to him (*I am going to conduct an experiment with my*

friends) now elude his understanding as he lacks complete consciousness of the operation's effects on himself. Thus, there is one detail in *our* Searle's thought experiment that demands attention: the symbols in his mind; even if they look like scribbles or seem incomprehensible, they do have some meaning, as they are part of the experiment using *h_Mori* and Mori's computer. The fact that he operates formal symbols does not mean he does not recognize that he exists within an experiment in a room shared by him and Mori. This illustrates the impossibility of isolating the symbolic dimension from the imaginary and the real ones. Searle is distinct from both a computer and *h_Mori*.

Both Turing and Searle are correct in their arguments (Turing closer to the imaginary and Searle closer to the symbolic), but one relies on the other's dimension and on *h_Mori* to function.

Turing's imaginary belief in *h_Mori* as a human is based on the symbolic operations in the written messages on the notes. Imaginary attribution to the humanlike image of messages only occurs when there is symbolic content. Searle's symbolic operations indicate automatism and rely on the imaginary belief that transforms the humanlike into the human itself. Symbolic repetition of codes in messages only occurs when there is imaginary belief. It is true, Mori is in the Symbolic and sharing the room with Searle. Searle and Mori each exemplify systems that divide the Symbolic ("two 'systems', both of which pass the Turing test" (Searle, 1980, p. 419))²⁴. However, Mori understands the messages, whereas Searle does not.

Everyone successfully passed their tests.

We are aware that *h_Mori*, created by Mori, works as a crowd-sourced medium for the "Turing test", Searle's "Japanese room", and Mori's "uncanny valley". Initially, all three eminent thinkers attribute *h_Mori* with a distinctly humanlike quality. By employing state-of-the-art technology, Mori enables *h_Mori* to play the role of a humanoid without triggering the "uncanny valley" response. Furthermore, the good design of *h_Mori* and robust programs created by Mori are necessary requirements for Searle to operate them, even indirectly. They understand *h_Mori* as a scientific way of testing due to its humanoid robot nature.

²⁴ We can classify the "two systems" as (1) Mori and (2) Searle/*h_Mori*.

Unexpectedly, however, after passing all the tests, h_Mori stops working. The intertwined dimensions of the imaginary and the symbolic lose their power to mediate communication when it stops working, leaving only its real carcass in view. According to Halfmann, irritation is all that is left. In this process, the imaginary of a humanlike entity loses consistency and what was assumed to be human (h_Mori) becomes a formless mass. The symbolic operations of formal symbols are now meaningless and, as a result, h_Mori only performs incorrect symbolic program repetitions. Within the Real, everything becomes an amorphous problem (there are issues, from wet floor to defective electricity and water pipes). Lacan's concept of the real, even if interpreted differently by us, is worth reflecting upon because it represents something unsymbolizable ("le réel en tant qu'impossible à dire" (Lacan, 1977, pp. 14-15)).

The failure of technological objects (like humanoid robots) to work represents the disruption of the reality mentioned by De Sousa Filho (2019, p. 58)²⁵ and the rupture of expectations mentioned by Halfmann (1996, p. 120)²⁶:

- 1) *The technological imaginary collapses.* Furthermore, the failure of technology to work represents a disruption in the communication operations.
- 2) *The technological symbolic collapses.* What is left? The *nicht-sinnhafte* part of technology without social function:
 - A) Technology as a *medium*: *nicht-sinnhafte* technology enables communication operations to be transmitted through itself.
 - B) Technology as an *installation*: *nicht-sinnhafte* technology does not work, it is a matter that requires observation.

Thus, a medium cannot be observed since it is employed, and an installation cannot be employed since it is observed (Halfmann, 1996, p. 116).

25 DeSousa Filho (2019, p. 58) states: The real "(...) causes and results in 'holes' in its appearance of coherence, consistency, accuracy, harmony, and abundance".

26 Another excerpt (pp. 151-52): "Ich erwarte, daß jeden Morgen mein Wagen anspringt; tut er es nicht, ist der alltägliche Kommunikationszusammenhang vielleicht schon gesprengt" (translated by us: "I expect my car to start every morning. If it does not, then maybe the everyday context of communication has already been shattered").

When h_Mori breaks down and stops mediating the communication, its operations, which were previously *nicht-sinnhaft*, also cease to work, requiring observation. Our three protagonists forcefully break down Real's door to observe what is "wrong". Nonetheless, they did not have time for full analysis and instead went to class and had a beer at the *Kneipe*. With h_Mori broken, the future experiments became inconclusive, unable to be confirmed or refuted. Once the humanoid was gone, the experiment context appeared as a real problem. The remaining data was a result of the scientific mediation called h_Mori.

Our humanoid robot h_Mori is an experimental tool that is effective when it mediates scientific information within the interdependent dimensions of imaginary, symbolic, and real. The installation of the real was achieved through symbolic operations, resulting in a surface-level image of a technologically available set. Thus, h_Mori is considered as a whole, an identity. But from h_Mori's perspective, the symbolic does not make sense. Its computer and microprocessor both do not make sense. Even its wires do not make sense. Metals, motors and sensors do not make sense. The logical rules of hardware computing exist only as a model that operates in the symbolic and imaginary dimensions, "taking advantage" of the real. In themselves, they do not make sense. The real remains indifferent to all of this. In terms of real, h_Mori is pure physical matter with cables and microcircuits through which electricity passes.

CONCLUSION: RETURNING TO THE WG

After having a beer, Turing, Mori, and Searle make their way back to the WG. It is time for everyone to return to their rooms. But they will have to leave their rooms at some point and fix the Real's situation. In this paper, I sought to mentally experience three dimensions contemplated by critical constructionism, which are relevant to three theses that employ a humanoid robot as a technological mediator. I placed Turing in the Imaginary, Mori and Searle in the Symbolic, and h_Mori in the Real. Within the Imaginary, Turing cannot assert that h_Mori is not Mori. In the Symbolic, Searle was an instantiation of h_Mori without any understanding, while Mori possessed complete understanding of the messages. Uncanny Valley

was avoided by *h_Mori* as a result of everyone playing their role flawlessly. *The reason all this happened is because h_Mori worked.* Everything fell apart when *h_Mori* stopped working: in the absence of a medium for experiments, there can be neither experiments nor *Gedankenexperimente*. Turing, Mori, and Searle must all depart from their respective rooms to observe the Real.

Due to the problematic and fragmented functioning of technological objects, it is difficult to distinguish a “cutting-edge technology” that does not blend with the concrete technologies that enable its operation. This disrupts the link between identity (such as the humanoid robot) and technological development. If we consider the “hierarchical” levels of technology and the added value in cutting-edge technologies, it may be worth maintaining a sense of relativism in the face of technological ideology. Through Franz Boas, Émile Durkheim, Marcel Mauss, Bronislaw Malinowski and others, anthropology has taught us that evaluations placing “advanced” cultures against “backward” cultures are rooted in prejudice (ethnocentrism). The differentiation of advanced and backward technologies undermines communication by favoring certain media over others, implying that some forms of communication are superior to others, while what truly matters is being together socially and emotionally.

REFERENCES

BOURDIEU, Pierre – *La Domination masculine*. Paris: Le Seuil, 1998. ISBN 9782020352512.

CASTORIADIS, Cornélius – *L'institution imaginaire de la société*. Paris: Le Seuil, 1975. ISBN 2020042525.

DE SOUSA FILHO, Alípio – *Revoke Ideology. Critical Constructionist Theory in the Human Sciences*. Oxford: Peter Lang, 2019. ISBN 9781789975482.

DURKHEIM, Émile – *Les formes élémentaires de la vie religieuse. Le système totémique en Australie*. Paris: PUF, 1990. ISBN 2130433383.

FREUD, Sigmund – *Der Wahn und die Traüme in W. Jensens ‚Gradiva‘*. Vienna: GW 7, 1999a. ISBN 3596503000.

FREUD, Sigmund – *Das Unbehagen in der Kultur*. Viena, GW 14, 1999b. ISBN 3596503000.

FREUD, Sigmund – *Civilization and its discontents*. London: SE 21, 1961. ISBN 9780701200671.

FRIEDKIN, William – *The Exorcist* [Film]. United States: Warner Bros, 1973.

HALFMANN, Jost – *Die gesellschaftliche „Natur“ der Technik. Eine soziologische Theorie der Technik*. Opladen: Leske + Budrich, 1996. ISBN 3810015504.

HERSEY, George – *Falling in Love with Statues. Artificial Humans from Pygmalion to the Present*. Chicago/London: University of Chicago Press, 2009. ISBN 9780226327792.

HOFFMANN, Ernst Theodor Amadeus – Der Sandmann. In HOFFMANN, Ernst Theodor Amadeus – *Sämtliche Werke*, 3. Frankfurt am Main: Deutscher Klassiker Verlag, 2004. ISBN 3618609051.

LACAN, Jacques – Le stade du miroir comme formateur de la fonction du je, telle qu'elle nous est révélée, dans l'expérience psychanalytique. *Revue Française de Psychanalyse*. Paris. 13/4 (1949) 449-455. ISSN 2105-2964.

LACAN, Jacques – *Écrits*. Paris: Seuil, 1966. ISBN 2020027526.

LACAN, Jacques – *Le Séminaire. Livre 1 Les écrits technique de Freud*. Paris: Seuil, 1975. ISBN 2020027682.

LACAN, Jacques – C'est à la lecture de Freud. In GEORGIN, Robert – *LACAN*. Cahiers Cistre, 3 (1977) 9-17. Lausanne, L'Âge d'Homme. ISBN 9789506021665.

LACAN, Jacques – *Le Séminaire. Livre 2 Le moi dans la théorie de Freud et dans la technique de la psychanalyse*. Paris: Seuil, 1978. ISBN 2020047276.

LACAN, Jacques – *Le Séminaire. Livre 3 Les psychoses*. Paris: Seuil, 1981. ISBN 2020060264.

LÉVI-STRAUSS, Claude – *Les Structures élémentaires de la parenté*. Paris: EHESS, 2017. ISBN 9782713227158.

LUHMANN, Niklas – *Soziale Systeme. Grundriß einer allgemeinen Theorie*. Frankfurt am Main: Suhrkamp, 1984. ISBN 9783518282663.

MORI, Masahiro – The uncanny valley. *IEEE Robotics & Automation Magazine*. 19:2 (2012) 98-100. ISSN 1558-223X.

PLATO – *The republic*. Cambridge: Cambridge University Press, 2000. ISBN 9780521484435.

228 What if Alan Turing, Masahiro Mori, and John Searle employed a humanoid robot in a *Gedankenexperiment*?

SEARLE, John – Minds, Brains and Programs. *Behavioral and Brain Sciences*. 3:3 (1980) 417–457. ISSN 0140-525X.

TURING, Alan – Computing Machinery and Intelligence. *Mind*. LIX:236 (1950) 433–460. ISSN 0026-4423.

WEBER, Max – Die „Objektivität“ sozialwissenschaftlicher und sozialpolitischer Erkenntnis. *Archiv für Sozialwissenschaft und Sozialpolitik*. 19:1 (1904) 22–87. ISSN 0174-819X.