

La proyección antropocéntrica del *Test de Turing* en la cultura de la IA

Anllely García Rojas^a

Resumen

El *Test de Turing* sigue desempeñando un papel filosófico importante (Moor 2001), especialmente en las investigaciones y reflexiones sobre la capacidad que posee una inteligencia artificial para engañar a las personas haciéndoles creer que es humano. Estas exploraciones pueden llegar a impactar significativamente en las narrativas y/o imaginarios colectivos, esto es, en la cultura científica en torno a la IA. De hecho, la posibilidad de que sistemas de IA pasen el *Test de Turing* en muchas ocasiones despierta un temor, pero también una profunda fascinación: el sueño de la conciencia artificial y la pesadilla de la rebelión de los robots ha estado presente en gran parte de la ciencia ficción, pero también en informes periodísticos y debates públicos (Natale 2021). Surge así, la pregunta que guiará el proyecto de ensayo: ¿En qué sentido el *Test de Turing* ha afectado la cultura en torno a la IA?

El presente ensayo pretende mostrar que el *Test de Turing* contiene la idea de que la IA reside no sólo en las técnicas de programación, sino también en las formas en que los humanos perciben e interactúan con el computador. El primer apartado expone la argumentación de Simone Natale (2021) en torno a la tesis de que el *Test de Turing* define la IA en términos de la perspectiva de los humanos, en sus interacciones con los computadores, otorgando un lugar destacado a la comunicación. El segundo apartado profundiza en el impacto del *Test de Turing* –y su sesgo antropocéntrico– en el contexto sociocultural, considerando también la perspectiva de Benjamin Bratton (2015), quien enuncia la necesidad de atender a la cultura de la IA o concepción popular de la IA, comprendida como la intersección entre la IA y la cultura humana.

^aUniversidad de Valparaíso, Valparaíso, Chile.
Contacto: anllelygarciarojas@gmail.com.

Se concluye que el *Test de Turing* proporciona claves interpretativas útiles para reflexionar acerca del impacto de la IA en las interacciones humano-computador.

Palabras clave: Inteligencia Artificial – Sesgo antropocéntrico – Cultura humana – Interacción humano-computador – Imaginarios colectivos.

Abstract

The Turing Test continues to play an important philosophical role (Moor 2001), especially in research and reflections on the ability of artificial intelligence to deceive people into believing it is human. These explorations can significantly impact narratives and/or collective imaginations, that is, the scientific culture surrounding AI. In fact, the possibility of AI systems passing the Turing Test often arouses fear, but also deep fascination: the dream of artificial consciousness and the nightmare of robot rebellion has been present in much of science fiction, but also in journalistic reports and public debates (Natale 2021). Thus, the question that will guide the essay project arises: In what sense has the Turing Test affected the culture surrounding AI?

This essay aims to show that the Turing Test contains the idea that AI resides not only in programming techniques, but also in the ways in which humans perceive and interact with the computer. The first section presents Simone Natale's (2021) argument around the thesis that the Turing Test defines AI in terms of the perspective of humans, in their interactions with computers, giving a prominent place to communication. The second section delves into the impact of the Turing Test –and its anthropocentric bias– in the sociocultural context, also considering the perspective of Benjamin Bratton (2015), who enunciates the need to attend to the culture of AI or popular conception of AI, understood as the intersection between AI and human culture. It is concluded that the Turing Test provides useful interpretive keys to reflect on the impact of AI on human-computer interactions.

Keywords: Artificial Intelligence – Anthropocentric bias – Human culture – Human-computer interaction – Collective imaginaries.

1. Introducción

El matemático británico Alan Turing (1912-1954) propuso el *Test de Turing* como reemplazo de la pregunta “¿Pueden pensar las máquinas?” en “Computing Machinery and Intelligence” (1950). Esto debido a la dificultad de hallar un consenso sobre el significado de términos como ‘máquina’, ‘pensar’ e ‘inteligencia’ y también a la imposibilidad de evaluar la experiencia o la conciencia de la máquina sin estar dentro o serlo. Un animado debate comenzó poco después de la publicación del artículo y continúa hoy, a través de críticas feroces, aprobaciones

entusiastas e interpretaciones contrastantes. Esto permite vislumbrar distintos hilos argumentativos que la propuesta de Turing abrió para el naciente campo de la IA, obligando a los lectores y comentaristas a considerar las posibilidades de inteligencia artificial en el ámbito formal académico y no exclusivamente desde la literatura de ciencia ficción (Natale 2021). Es relevante mencionar que las contribuciones de Turing fueron decisivas para el posterior surgimiento de la investigación en IA. Influyó directamente en las teorías de los matemáticos John McCarthy y Marvin Minsky, quienes en 1956 organizaron la conferencia “Dartmouth Research Project on Artificial Intelligence”, donde se llegó a un consenso respecto al nombramiento e inauguración de la disciplina.

Actualmente, el *Test de Turing* sigue desempeñando un papel filosófico importante (Moor 2001), especialmente en las investigaciones y reflexiones sobre la capacidad que posee la inteligencia artificial para engañar a las personas haciéndoles creer que es humana. Estas exploraciones pueden llegar a impactar significativamente en las narrativas y/o imaginarios colectivos, esto es, en la cultura o concepción popular de la IA. De hecho, la posibilidad de que sistemas de IA pasen el *Test de Turing* en muchas ocasiones despierta temor, pero también profunda fascinación: el sueño de la conciencia artificial y la pesadilla de la rebelión de los robots ha estado presente en gran parte de la ciencia ficción, pero también en informes periodísticos y debates públicos (Natale 2021). Surge así, la pregunta que guiará el proyecto de ensayo: ¿En qué sentido el *Test de Turing* ha afectado la cultura en torno a la IA?

El presente ensayo pretende mostrar que el *Test de Turing* contiene la idea de que la IA reside no sólo en las técnicas de programación, sino también en las formas en que los humanos perciben e interactúan con el computador. El primer apartado expone la argumentación de Simone Natale (2021) en torno a la tesis de que el *Test de Turing* define la IA en términos de la perspectiva de los humanos, en sus interacciones con los computadores, otorgando un lugar destacado a la comunicación. El segundo apartado profundiza en el impacto del *Test de Turing* –y su sesgo antropocéntrico– en el contexto sociocultural, considerando también la perspectiva de Benjamin Bratton (2015), quien enuncia la necesidad de atender a la cultura de la IA o concepción popular de la IA, comprendida como la intersección entre la IA y la cultura humana. Se concluye que el *Test de Turing* proporciona claves interpretativas útiles para reflexionar acerca del impacto de la IA en las interacciones humano-computador.

2. La dimensión comunicacional del *Test de Turing*

En 1950, Alan Turing publicó su famoso artículo “Computing Machinery and Intelligence”, en la revista *Mind*. Una de las contribuciones más significativas para las investigaciones de las principales ramificaciones de la Ingeniería Informática; la teoría de la informática, el diseño de hardware-software y el estudio de la inteligencia artificial (Shieber 2004). En la propuesta teórica, Turing describió lo que actualmente llamamos *Test de Turing*, y al que se refirió mediante el Juego de la Imitación, del cual existen diversas versiones. No obstante, el presente proyecto refiere a la interpretación estándar, comprendida como un método empírico que evalúa si una máquina –un computador digital– puede exhibir un comportamiento inteligente.

El *Test de Turing* involucra a tres jugadores que participan en actos de comunicación: el computador digital, un jugador humano y un interrogador o juez humano. Este último tiene el desafío de distinguir cuál de los otros dos jugadores es la persona y cuál es la máquina programada a través de mensajes escritos –Turing imaginó llevar a cabo estas conversaciones por medio de una teleimpresora– (Copeland et al 2017). Los tres participantes se mantienen estrictamente fuera de contacto entre sí, sin espiar. El interrogador hace preguntas tan penetrantes y amplias como guste, y al computador digital se le permite hacer todo lo posible para forzar una identificación incorrecta. La máquina usará “todo tipo de trucos para parecerse más a un hombre”, expresó Turing en la discusión que mantuvo con Newman, R.B Braithwaite y G. Jefferson, la cual fue grabada y transmitida por BBC radio en 1952. Turing precisa respecto a estas temáticas de la siguiente forma:

BRAITHWAITE: Would the questions have to be sums, or could I ask it what it had had for breakfast?

TURING: Oh yes, anything. And the questions don't really have to be questions, any more than questions in a law court are really questions. You know the sort of thing. “I put it to you that you are only pretending to be a man” would be quite in order. Likewise the machine would be permitted all sorts of tricks so as to appear more man-like, such as waiting a bit before giving the answer, or making spelling mistakes, but it can't make smudges on the paper, any more than one can send smudges by telegraph. We had better suppose that each jury has to judge quite a number of times, and that sometimes they really are dealing with a man and not

a machine. That will prevent them saying “It must be a machine” every time without proper consideration. (Shieber 2004, pp. 118-119)

A continuación, se expone el ejemplo de Turing del tipo de conversación que podría ocurrir entre un juez y una computadora que evade con éxito la identificación:

JUDGE: In the first line of your sonnet which reads “Shall I compare thee to a summer’s day”, would not “a spring day” do as well or better?

MACHINE: It wouldn’t scan.

JUDGE: How about “a winter’s day”? That would scan all right.

MACHINE: Yes, but nobody wants to be compared to a winter’s day.

JUDGE: Would you say Mr Pickwick reminded you of Christmas?

MACHINE: In a way.

JUDGE: Yet Christmas is a winter’s day, and I do not think Mr Pickwick would mind the comparison.

MACHINE: I don’t think you’re serious. By a winter’s day one means a typical winter’s day, rather than a special one like Christmas. (Copeland et al 2017, p. 270)

El propósito del computador digital es engañar al interrogador haciéndole creer que es un humano mediante la imitación del comportamiento verbal. Mientras que el otro participante busca ayudar al interrogador revelando su identidad humana. Si el interrogador no logra hacer la correcta identificación entre el jugador humano y el computador digital, entonces este último supera la prueba. De este modo, el objetivo del *Test de Turing* es verificar si existe inteligencia mediante la imitación eficaz (González 2007). Tras definir el Juego de la Imitación, la nueva agenda a discutir se trasladó de “¿Pueden pensar las máquinas?” a la pregunta “¿Existen computadoras digitales imaginables que tendrían un buen desempeño en el juego de la imitación?” (Turing 1950, p. 442).

El artículo de Turing fue ambiguo en muchos aspectos, lo cual favoreció el surgimiento de diferentes puntos de vista y controversias sobre

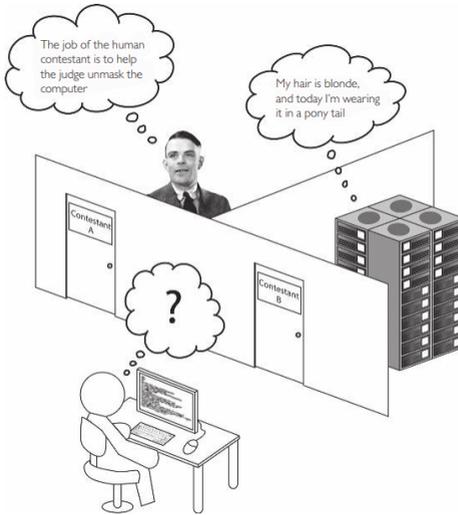


Figura 2: El *Test de Turing*. El juez debe decidir qué concursante es el computador (Copeland et al 2017).

el significado del *Test de Turing*. Sin embargo, una de las implicaciones clave para el campo de la IA es evidente. La pregunta, dice Turing a sus lectores, no es si las máquinas son o no capaces de pensar. Es, en cambio, si creemos que las máquinas son capaces de pensar, en otras palabras, si estamos preparados para aceptar el comportamiento de las máquinas como inteligente (Natale 2021, p. 20).

Simone Natale, doctor en estudios de Comunicación y Medios, sostiene en *Deceitful Media: Artificial Intelligence and Social Life After the Turing Test* que la capacidad del computador para imitar la forma en que se comportan los humanos en una conversación es obviamente uno de los factores para superar el *Test de Turing*, pero también lo es la participación activa de los jugadores humanos. Tanto el criterio utilizado por el interrogador o juez como las motivaciones personales del jugador humano que interviene en la conversación son variables que pueden influir el resultado de la prueba (Natale 2021, pp. 20-21).

Things such as their backgrounds, biases, characters, genders, and political opinions will have a role in both the decisions of the interrogator and the behavior of the human

agents with which the interrogator interacts. A computer scientist with knowledge and experience on AI, for instance, will be in a different position from someone who has limited insight into the topic. Likewise, human players who act as conversation partners in the test will have their own motivations and ideas on how to participate in the test. Some, for instance, could be fascinated by the possibility of being exchanged for a computer and therefore tempted to create ambiguity about their identities. Because of the role human actors play in the Turing test, all these are variables that may inform its outcome (Natale 2021, pp. 20-21).

Jennifer Rhee señala perspicazmente en “La promesa de la identificación errónea” que en el *Test de Turing* la responsabilidad del éxito o falla no recae únicamente en las habilidades de la máquina, sino que en gran medida recae en el ser humano (Rhee 2010). A partir de esto, es posible afirmar que los elementos decisivos para el resultado de la prueba comprenden un sesgo antropocéntrico, arraigado en la comprensión de la noción de inteligencia, más bien en la expectativa y percepción de un comportamiento inteligente similar al humano. En este sentido, la decisión de Turing respecto al diseño de la prueba, particularmente del papel del interrogador humano, se basa en un sesgo antropocéntrico. Este sesgo humano se evidencia al considerar que “(...) Turing no considera el reemplazo de los interrogadores, o que haya un programa que desempeñe la misma función que los mismos” (González 2016).

Para ilustrar la magnitud del impacto que tiene este sesgo en el resultado del *Test de Turing*, se presenta el caso de los resultados del “Premio Loebner” en el año 2000, un concurso anual en que diversos programas de computadores se involucran en una versión de la prueba:

In the 2000 contest, for example, members of the jury judged a human as a machine ten times, but did not judge any machine as a human being; and in the 2003 contest they judged a human as “definitely a machine” four times, but did not judge any computer as “definitely a human” (Copeland et al 2017, p. 290)

Al enfocarse no en el ser humano, no en la máquina, sino en el umbral y/o nexos entre estos dos agentes, es posible interpretar nuevas significaciones relacionales, no a través de definiciones estáticas e independientes, sino más bien a través de las interacciones y encuentros

entre humanos y máquinas (Rhee 2010). La lectura de Natale sostiene que, al definir la inteligencia de una máquina en términos de la capacidad de un computador para superar la prueba, Turing incluyó a los humanos en la fórmula, haciendo que sus ideas y sesgos, así como su psicología y carácter, sean una variable crucial en la construcción de máquinas “inteligentes”.

En otras palabras, independientemente de las intenciones iniciales de Turing, el *Test de Turing* ofrece definir la inteligencia artificial en términos de la perspectiva humana en su interacción con los computadores digitales. Al ser así, la centralidad de los seres humanos sugiere que la inteligencia artificial reside no sólo en circuitos y técnicas de programación, sino también en las formas en que los humanos perciben e interactúan con las máquinas. Esta interpretación, según Natale, estimuló a investigadores y desarrolladores de tecnologías a modelar comportamientos humanos y estados de ánimo con el fin de diseñar sistemas de IA interactivos más efectivos (Natale 2021).

The uncertainty that derives from this is often indicated to be one of the test’s shortcomings. The test appears entirely logical and justified, however, if one sees it not as an evaluation of the existence of thinking machines but as a measure of humans’ reactions to communications with machines that exhibit intelligent behavior. From this point of view, the test made clear for the first time, decades before the emergence of online communities, social media bots, and voice assistants, that AI not only is a matter of computer power and programming techniques but also resides—and perhaps especially—in the perceptions and patterns of interaction through which humans engage with computers (Natale 2021, p. 21).

Al situar a los humanos en el centro del diseño del *Test de Turing*, el teórico británico proporcionó un complejo escenario que surge de las interacciones humano-computador, situando a la comunicación—y una forma particular de interacción social engañosa—como un factor decisivo. Así, las tecnologías de IA contemporáneas podrían ser concebidas en términos de su credibilidad para los usuarios humanos, desde su dimensión relacional que emerge especialmente dentro de la interacción entre humanos y máquinas. Brian Christian, en un reflexivo libro en torno a su experiencia como participante en el “Premio Loebner”, destaca el aspecto comunicacional del *Test de Turing*:

Here's the thing: beyond its use as a technological benchmark, beyond even the philosophical, biological, and moral questions it poses, the Turing test is, at bottom, about the act of communication. I see its deepest questions as practical ones: How do we connect meaningfully with each other, as meaningfully as possible, within the limits of language and time? How does empathy work? What is the process by which someone comes into our life and comes to mean something to us? These, to me, are the test's most central questions—the most central questions of being human (Christian 2011, p. 9).

Natale sostiene que el diseño del *Test de Turing* se basa en las interacciones entre humanos y computadores, destacando que Turing experimentó la urgencia de incluir detalles precisos acerca del medio de comunicación; justamente con la finalidad de garantizar la validez de la prueba, el interrogador requiere comunicarse con el jugador humano y el computador sin recibir ninguna pista sobre sus identidades, excepto el contenido de los mensajes escritos. En este sentido, la comunicación entre los humanos y el computador en el *Test de Turing* estaba destinada a tener un carácter anónimo e incorpóreo (Natale 2021). En palabras de Turing:

In order that tones of voice may not help the interrogator the answers should be written, or better still, typewritten. The ideal arrangement is to have a teleprinter communicating between the two rooms. Alternatively the question and answers can be repeated by an intermediary (Turing 1950, p. 434).

La mediación tecnológica del *Test de Turing*, a saber, la teleimpresora, aspira a lograr que tanto los participantes digitales como los humanos participen en la comunicación como información pura, es decir, como un conjunto de datos independiente del autor. Esto sugiere que la solución dada por Turing implica una comprensión aguda del papel de los medios de comunicación. En la actualidad, el uso de diferentes tecnologías e interfaces para enviar mensajes evidencia tanto la naturaleza computacional de esos programas como el efecto de su comunicación, es decir, cómo impactan en los usuarios humanos (Natale 2021). Aquí un ejemplo, utilizado por Natale, que realiza el impacto de los medios e interfaces en los efectos de la comunicación:

Similarly, where voice is concerned, the nature of communication will be different if AI voice assistants are applied to domestic environments, like Alexa, embedded in smartphones and other mobile devices, like Siri, or employed to conduct telephonic conversations, as in automated customer services or Google's in-progress Duplex project (Natale 2021, p. 24).

Desde la perspectiva de Natale, el *Test de Turing* no solo propone definir la inteligencia artificial en términos de la perspectiva humana, esto es, desde su dimensión relacional, sino también puede ser comprendida como medio que posibilita y regula dinámicas específicas de comunicación entre humanos y computadores. Esto es de especial importancia, ya que la computación no siempre se ha visto vinculada con la comunicación.

En 1950, cuando Turing investigaba, las formas de interacción entre los usuarios humanos y las computadoras se discutían mínimamente en tanto los computadores eran considerados principalmente como herramientas de cálculo. Así, la propuesta de Turing implicó un salto visionario en la historia de la computación y la IA que contempló un cambio significativo en las actitudes culturales (Natale 2021). Y es que el matemático británico no solo comentó sobre el desarrollo de tecnologías más funcionales, sino que también estimó un cambio cultural al expresar:

The original question, "Can machines think?" I believe to be too meaningless to deserve discussion. Nevertheless I believe that at the end of the century the use of words and general educated opinion will have altered so much that one will be able to speak of machines thinking without expecting to be contradicted. I believe further that no useful purpose is served by concealing these beliefs (Turing 1950, p. 442).

Turing tenía razón al darse cuenta de que las actitudes culturales cambiarían, como consecuencia tanto de la evolución de la informática como de las experiencias de las personas con las tecnologías. De esta manera, el *Test de Turing* define la IA en términos de la perspectiva de los humanos, en sus interacciones con los computadores, otorgando un lugar destacado a la comunicación.

3. El impacto del *Test de Turing* en la cultura de la IA

El *Test de Turing* estimuló animadas discusiones teóricas en áreas como la psicología, la informática y la filosofía de la mente y, por tanto, resultó ser una excelente provocación para reconocer el potencial de la informática. Como consecuencia, a principios de la década de 1990, Hugh Loebner y Robert Epstein decidieron organizar un concurso público real basado en la propuesta de Turing, al que denominaron *Premio Loebner*. Desde la perspectiva de Natale, los factores que impulsaron dicho concurso no fueron exclusivamente técnicos, sino también culturales; el desarrollo de tecnologías digitales y el Internet habían creado condiciones ideales para llevar a cabo el *Premio Loebner*, por lo que, es ante todo una cuestión de comunicación y entretenimiento; su intención principal es atraer al público. Al ser así, se puede esperar que este evento y, en general, los concursos humanos versus máquinas tendrían un impacto positivo en la cultura de la IA o concepción popular de la IA, nociones que refieren a las formas en que la IA influye en la cultura humana y viceversa.

Natale considera que la popularidad de tales eventos atestigua un temor, pero también una profunda fascinación en las culturas contemporáneas, la cual es evidente en las representaciones ficticias de la IA. El sueño de la conciencia artificial y la pesadilla de la rebelión de los robots ha estado presente en gran parte de la ciencia ficción, pero también en informes periodísticos, debates públicos y/o en actuaciones públicas, donde la IA se celebra como una forma de espectáculo, como el concurso del *Premio Loebner* (Natale 2021). La difusión de estos imaginarios puede aumentar la posibilidad de que sistemas informáticos relativamente simples se perciban como inteligentes, exagerando los logros del campo de la IA. Respecto al lugar del engaño en la interacción humano-computador y las disposiciones anímicas que incentiva, Natale expresa:

People seem to have a somehow ambivalent, yet deep, desire to see the myth of the thinking machine come true, which shapes many journalistic reports and public debates about the Turing test (...) the Loebner Prize competition says more about how people are deceived than about the current state of computing and AI. The Loebner Prize competition has become an arena of deception that places humans' liability for being deceived at center stage. Yet this does not make this contest any less interesting and useful for unders-

tanding the dynamics of communicative AI systems (Natale 2021, p. 104).

El *Test de Turing* sugiere evaluar la inteligencia del computador considerando su capacidad para simular la forma en que se comportan los humanos en una conversación mediante el engaño. La idea de que la capacidad de mentir es una característica definitoria de los humanos se encuentra en la base de la prueba. En este sentido, incluir la mentira y el engaño en la definición de la IA vuelve la atención a la centralidad de los humanos y, por tanto, al sesgo antropocéntrico y/o lo que Benjamin Bratton en *Outing Artificial Intelligence: Reckoning with Turing Tests* denominará “falacia antropocéntrica”; la creencia de que toda inteligencia debe ser similar o idéntica a la humana. Esto llevaría a pensar la IA exclusivamente en términos de la perspectiva humana.

Airplanes do not fly like birds fly, and we certainly do not try to trick birds into thinking that airplanes are birds in order to test whether those planes “really” are flying machines. Why do it for AI then? Today the vast majority of core AI research is not focusing Turing Test as anything like a central criterion of success, and yet in our general discourse about AI, the test’s anthropocentrism still holds such conceptual importance (Bratton 2015, p. 74).

La centralidad de la perspectiva humana en la definición de la IA puede ser interpretada como una expresión del antropocentrismo, el cual tiene consecuencias directas en las dinámicas humano-computador y, por consiguiente, en un contexto social y cultural. El *Test de Turing* estableció el engaño como un efecto probable de dichas interacciones, ya que el computador “ganará” el juego si es capaz de engañar al interrogador humano, es decir, hacerle creer que el computador digital es humano mediante el comportamiento verbal. En este sentido, puede ser interpretada como una prueba de detección de mentiras en la que no es posible que el interrogador confíe en sus compañeros de conversación, pues todo mensaje emitido por la máquina es falso (Natale 2021).

Es por este motivo, quizá, que la posibilidad de que los sistemas de IA superen el *Test de Turing* a menudo despierta temores; ya que las máquinas engañarán a los humanos con el fin de avanzar un paso más en su rebelión contra sus creadores (Natale 2021). De este modo, el sesgo o falacia antropocéntrica, presente en el *Test de Turing*, se proyecta en la cultura o concepción popular de la IA, reflejándose en innumerables

narrativas o ficciones que otorgan características humanas, como la dimensión afectiva o la intencionalidad –ira, celos, confusión, avaricia, orgullo, deseo, entre otros– a los sistemas de IA (Bratton 2015). Esto se encuentra dramatizado en innumerables narrativas de ciencia ficción, desde “Blade Runner” de Ridley Scott hasta “Her” de Spike Jonze.

La falacia antropocéntrica configura un prisma a través del cual gran parte de la cultura humana comprende las interacciones con la IA, por lo que Bratton recomienda un análisis crítico de sus implicaciones a nivel social, cultural y político. Debido a que, si se busca la definición de IA de manera equivocada, esta puede surgir en formas que son innecesariamente difíciles de reconocer, amplificando sus riesgos y la opacidad o mistificación del funcionamiento de la computación y, a su vez, de los logros del campo de la IA (Bratton 2015).

We need a popular culture of A.I. that is less parochial and narcissistic, one that is based on more than simply looking for a machine version of our own reflection. As a basis for staging encounters between various A.I.s and humans, that would be a deeply flawed precondition for communication. Needless to say, our historical track record with “first contacts,” even among ourselves, does not provide clear comfort that we are well-prepared (Bratton 2015).

La relación entre la IA y el engaño configura una problemática relevante para las interacciones entre humanos y computadores que suscita profundas reflexiones. Por este motivo, la preocupación de Natale se centra en el lugar del engaño, defendiendo que “la IA debería situarse dentro de la trayectoria más larga de los medios engañosos que incorporan engaño banal en su función” (Natale 2021, p. 31). El *Test de Turing* contribuye a la justificación de dicha afirmación al momento de considerar el componente lúdico del engaño administrado por el computador.

Se suele atribuir una connotación negativa al engaño, pero su integración en actividades lúdicas es un recordatorio de que las personas buscan activamente situaciones en las que pueden ser engañadas, siguiendo un deseo o necesidad compartida (Natale 2021). En este sentido, el *Test de Turing* presenta la posibilidad de que los humanos puedan ser engañados por las computadoras en un contexto lúdico y voluntario. A continuación, se expone el ejemplo de Natale que muestra que cierto grado de engaño es inofensivo e incluso funcional en una interacción cotidiana:

Alexa and Siri are perfect examples of how this works in practice: the use of human voices and names with which these “assistants” can be summoned and the consistency of their behaviors stimulate users to assign a certain personality to them. This, in turn, helps users to introduce these systems more easily into their everyday lives and domestic spaces, making them less threatening and more familiar. Voice assistants function most effectively when a form of playful and willing deception is embedded in the interaction (Natale 2021, pp. 29-30).

La proyección de la falacia antropocéntrica, materializada en el mito tecnológico que se centra exclusivamente en la posibilidad de crear máquinas pensantes en un nivel técnico, aleja la atención del hecho de que las implicaciones más significativas de los sistemas de IA no se encuentran necesariamente en un futuro lejano, sino que en las interacciones cotidianas. Esto es de suma importancia si consideramos que, en su día a día, los humanos deciden construir un espacio de interacción recíproca con diversos sistemas de IA que modifican el mundo que los circunda. Por este motivo, Natale enuncia la urgencia de reflexionar sobre el lugar del engaño interdisciplinariamente; en la encrucijada entre las ciencias de la computación, las ciencias económicas, las ciencias cognitivas, las ciencias sociales y las humanidades.

4. Conclusión

Un animado debate comenzó poco después de la publicación del artículo “Computing Machinery and Intelligence” (1950), con fuertes críticas, apoyo entusiasta e interpretaciones contrastantes, y continúa hasta el día de hoy. Como se demostró, el *Test de Turing* sigue desempeñando un papel filosófico importante, especialmente en las reflexiones sobre la dimensión relacional de la IA, la cual emerge especialmente dentro de la interacción humano-computador. De esta manera, aun si el *Test de Turing* no refleja el funcionamiento de los sistemas modernos de la IA, proporciona claves interpretativas útiles para comprender las comunicaciones entre humanos y máquinas, situadas en un contexto social y cultural. La influencia de la inteligencia artificial en los comportamientos sociales evidencia la relevancia de la temática desarrollada en este ensayo.

Siguiendo las reflexiones de Benjamin H. Bratton (2015) y Simone Natale (2021), la argumentación desplegada ante la pregunta: ¿En qué

sentido el *Test de Turing* ha afectado la cultura en torno a la IA? Se centró en mostrar, en el primer apartado, que independientemente de las intenciones iniciales de Turing, el *Test de Turing* contiene la idea de que la IA reside no sólo en las técnicas de programación, sino también en las formas en que los humanos perciben e interactúan con el computador. Así, la definición de inteligencia artificial depende de la percepción humana más que de alguna característica específica de las máquinas. De este modo, el *Test de Turing* no solo reconoce la dimensión relacional de la IA, sino también permite interpretarla como medio que posibilita y regula dinámicas específicas de comunicación entre humanos y computadores. El hecho de que el *Test de Turing* planteó una situación en la que un interrogador humano era propenso al engaño por la máquina suscitó una profunda controversia en el campo de la IA desde sus inicios. Por esto, en el segundo apartado se profundizó en el impacto del *Test de Turing*, particularmente desde la falacia antropocéntrica, en el contexto sociocultural, exponiendo que también se debe considerar el hecho de que la prueba se enmarca en lo lúdico y lo inofensivo de este engaño. De esta forma, aquello que se consideraba un simple objeto de la ciencia parece convertirse en un verdadero sujeto, ya que presenta la capacidad de tejer una red de relaciones con los humanos. Con todo, el presente escrito vislumbra en el *Test de Turing* un espacio de intersección entre la IA y la cultura humana, brindando un punto común para el desarrollo tecnológico de sistemas IA y los medios de comunicación o interfaces que posibilitan una reflexión más amplia de las implicaciones epistemológicas de la interacción humano-computador. Es necesario ahondar en estas temáticas desde la interdisciplinariedad, para enfrentar los desafíos éticos y políticos que surgen en este contexto.

Referencias

- Bernhardt, C. (2016). *Turing's vision: the birth of computer science*. Cambridge, Massachusetts: The MIT Press.
- Bratton, B. (2015). Outing Artificial Intelligence: Reckoning with Turing Tests, en Pasquinelli, M. *Alleys of Your Mind. Augmented Intelligence and Its Traumas*. Lüneburg: meson press , pp. 69–80.
- Bratton, B. (2015, febrero 23). Outing A.I.: Beyond the Turing Test. *The New York Times*. Recuperado de <https://>

//archive.nytimes.com/opinionator.blogs.nytimes.com/2015/02/23/outing-a-i-beyond-the-turing-test/.

Copeland, J. (2012). *Turing: Pioneer of the Information Age*. United Kingdom: Oxford University Press.

Copeland, J., Bowen, J., Sprevak, M., & Wilson, R. (2017). *The Turing Guide*. New York: Oxford University Press.

Copeland, J. (Ed.). (2004). *The essential Turing: Seminal Writings in Computing, Logic, Philosophy, Artificial Intelligence, and Artificial Life, plus The Secret of Enigma*. Oxford: Oxford University Press.

Christian, B. (2011). *The Most Human Human: What Talking with Computers Teaches Us About What It Means to Be Alive*. New York: Doubleday.

González, R. (2007). El Test de Turing: dos mitos, un dogma. *Revista de Filosofía*, 63 pp. 37–53.

González, R. (2016). Dos criterios para la presencia de estados mentales: Descartes y Turing. *Cinta Moebio*, 56 , pp. 159–171.

Gunkel, D. (2018). *Gaming The System: Deconstructing Video Games, Games Studies, and Virtual Worlds*. Bloomington: Indiana University Press.

Moor, J. (Ed.). (2003). *The Turing Test: The Elusive Standard of Artificial Intelligence*. Nueva York: Springer Science+Business Media.

Natale, S. (2021). *Deceitful media: artificial intelligence and social life after the Turing test*. New York: Oxford University Press.

Oppy, G., & Dowe, D. (2003). The Turing Test. *The Stanford Encyclopedia of Philosophy*. Recuperado de <https://plato.stanford.edu/archives/win2021/entries/turing-test/>.

Turing, A. (1950). Computing machinery and intelligence (Trad. Fuentes 2010). *Mind*, Vol. 59, ed. 236 pp. 433–460. doi:10.1093/mind/lix.236.433

Rhee, J. (2010). Anthropomorphic Attachments in U.S. Literature, Robotics, and Artificial Intelligence. Duke University.

Shieber, M. (Ed.) (2004). *The Turing test: verbal behavior as the hallmark of intelligence*. Cambridge, Massachusetts: The MIT Press.