

El enigma ontoepistemológico de la consciencia. Acerca del transhumanismo y la analogía computacional del cerebro*

*The Onto-Epistemological Enigma of Consciousness. On Transhumanism
and the Computational Analogy of the Brain*

 Carlos Hugo Sierra**



* Este artículo constituye un esbozo preliminar de un estudio de mayor alcance en torno al trasfondo epistémico que subyace en ciertos asuntos nucleares de la neurociencia contemporánea, desarrollado en el marco de la docencia sobre neurofilosofía, correspondiente al Diplomado Avanzado de Derecho, Tecnología e Innovación de la Universidad de Antioquia.

** Universidad de Antioquia, Medellín, Colombia. Correo electrónico: chugo.sierra@udea.edu.co

Fecha de recepción: 13 de julio de 2022

Fecha de aceptación: 4 de octubre de 2022

Cómo referenciar / How to cite

Sierra, C. H. (2023). El enigma ontoepistemológico de la consciencia. Acerca del transhumanismo y la analogía computacional del cerebro. *Trilogía Ciencia Tecnología Sociedad*, 15(29), e2458. <https://doi.org/10.22430/21457778.2458>

Resumen: el propósito principal de este artículo es presentar un análisis crítico sobre algunos de los planteamientos epistemológicos y filosóficos más destacados que se han elaborado desde el transhumanismo contemporáneo en relación con la actividad anatómico-funcional del cerebro y a la naturaleza de la consciencia. Desde ese punto de vista, se pretende considerar si el modelo hermenéutico tecnológico-instrumental que se promueve desde los planteamientos de la corriente transhumanista supone una variante comprensiva alternativa a la hora de dilucidar la complejidad ontoepistemológica de la consciencia. Para ello, se lleva a cabo una exploración sobre el modo en que el transhumanismo se ha centrado en ciertos campos tecnocientíficos para desarrollar una prospectiva concreta asociada a las potencialidades cognitivas del ser humano. En ese sentido, se toma en consideración ciertos avances concretos de la neurotecnología de vanguardia y de la neuromodulación artificial (implantología neural, dispositivos neuromórficos, reprogramación neuronal, cultivo controlado en laboratorio de cerebroides, etc.), así como de la inteligencia artificial. De este análisis se desprende un profundo escepticismo en torno a la posibilidad de que el enfoque transhumanista ofrezca un modelo prometedor para abordar exitosamente el antiguo problema de la relación mente-cerebro, ya que, pese al atractivo de algunas de sus especulaciones más audaces en torno a la condición bioantropológica del hombre y de la ontología de lo real, ha optado por actualizar una visión reduccionista y fisicalista de la mente humana.

Palabras clave: neuromodulación artificial, ontología digital, relación mente-cerebro, teoría computacional de la mente, transhumanismo.

Abstract: The main purpose of this reflection paper is to present a critical analysis of some of the most prominent epistemological and philosophical ideas developed by contemporary transhumanism in relation to the anatomical-functional activity of the brain and the nature of consciousness. From this point of view, this paper considers whether the technological-instrumental hermeneutic model promoted by the transhumanist movement is an alternative, comprehensive variant to elucidate the onto-epistemological complexity of consciousness. To this end, this paper explores the way in which transhumanism has focused on certain techno-scientific fields in order to develop a concrete prospective associated with the cognitive potentialities of the human being. In this sense, it discusses certain specific advances in cutting-edge neurotechnology and artificial neuromodulation (neural implants, neuromorphic devices, neuronal reprogramming, controlled laboratory culture of cerebroids, etc.), as well as artificial intelligence. This analysis results in deep scepticism as to whether the transhumanist approach offers a promising model for successfully tackling the age-old problem of the mind-brain relationship because, despite the appeal of some of its boldest speculations about the bio-anthropological condition of man and the ontology of the real, it has chosen to update a reductionist and physicalist view of the human mind.

Keywords: Artificial neuromodulation, digital ontology, mind-brain relationship, computational theory of mind, transhumanism.

«No sé quién soy. No soy lo que sé»
Angelus Silesius (Johannes Scheffler)
El peregrino querúbico (1657)

INTRODUCCIÓN

El 14 de agosto de 1872, el célebre médico y fisiólogo alemán, Emil Du Bois-Reymond (1818-1896)¹, intervino en una de las sesiones plenarias del congreso de científicos y médicos alemanes en la ciudad de Leipzig con una breve exposición, titulada *Sobre los límites del conocimiento de la naturaleza* (Du Bois-Reymond, 1872) que, sin sospecharlo, iba a levantar una gran polvareda entre los círculos científicos europeos. En el fondo no era para menos. Uno de los más destacados cabecillas de la corriente fisiológica antivitalista de Berlín (junto con Hermann Helmholtz, Ernst Brücke y Carl Ludwig), que ya desde 1847 alentaba con rabiosa determinación la irrupción del mecanicismo fisicoquímico en las ciencias de la vida, el brillante impulsor de la neurociencia moderna a través de sus asombrosos trabajos experimentales en el campo de la electrofisiología², mostraba, para sorpresa y conmoción del entregado auditorio, un sombrío escepticismo respecto a las esperanzas que aquel siglo había depositado en la empresa científica. La ciencia, y más en concreto la perspectiva estrictamente fiscalista, se topaba en su apoteósico avance con determinados límites epistemológicos que se tornaban verdaderamente infranqueables cuando uno se enfrentaba a ciertos enigmas de la realidad. Uno de ellos, que es el que nos atañe en este escrito, tenía que ver con la propia naturaleza de la consciencia. Y es que Du Bois-Reymond se mostraba plenamente convencido de que, a lo largo de la historia, ninguna perspectiva filosófica, por muy sofisticado que fuera su edificio teórico, había logrado explicar de manera satisfactoria el modo en que la mente interactúa con el cuerpo. No valía la pena, por tanto, malgastar excesivo tiempo en explorar los entresijos de la metafísica dualista cartesiana, la propuesta ocasionalista ideada por Nicolas Malebranche o la doctrina leibniziana sobre una quimérica armonía preestablecida. Tan solo el monismo fiscalista que él propugnaba se había acercado con un mayor grado de precisión a este problema inveterado, en la medida en que parecía evidente que la consciencia no era sino un efecto evolutivo de la materia, una especie de epifenómeno resultante del desarrollo estructural del sistema nervioso.

No obstante, esta peculiar supeditación ontológica (o sea, la de la mente respecto al cerebro), cuyo meollo quedaba toscamente ilustrado con la famosa expresión del zoólogo y filósofo Vogt (1874) de que «el pensamiento es al cerebro lo que la bilis es al hígado o la orina a los riñones» (p. 323), no acababa de despejar en absoluto las dudas de Du Bois-Reymond, ya que, en última instancia, reconocía (muy a su pesar) que, aun cuando en algún momento se obtuviese un perfecto conocimiento del cerebro, este no iba a proporcionar ninguna pista concluyente para discernir la naturaleza de la experiencia humana.

¹ Gabriel Finkelstein es, sin duda, uno de los historiadores que con mayor cuidado ha explorado la fascinante aventura vital y académica de Emil Du Bois-Reymond. Véase al respecto, Finkelstein (2013).

² Con la introducción de avances en los galvanómetros y la invención del potenciómetro o del magnetoeléctrometro, Du Bois-Reymond logra captar, mediante estimulación farádica, ondas de variación eléctrica en los músculos y fibras nerviosas de animales y, posteriormente, en los seres humanos.

En definitiva, nos hallábamos ante uno de los escasos aspectos de la naturaleza que no podían ser reducidos a un substrato material. Siendo así, el determinismo mecanicista, tal como lo llegó a concebir Du Bois-Reymond, resultaba también insuficiente para servir de puente entre el universo mental y la actividad cerebral. De esta forma, su agnosticismo científico quedaba resumido emblemáticamente para la posteridad en la archiconocida máxima *Ignoramus et ignorabimus* («desconocemos y desconoceremos») y, con ello, se daba el pistoletazo de salida a una apasionada polémica sobre los enigmas del universo y la capacidad de la ciencia para dilucidarlos (*Ignorabimus-Streit*) que involucraría a gran parte de la élite de la academia científica europea del momento: Rudolf Virchow, James Clerk Maxwell, Karl Nägeli, Élie de Cyon, Ernst Haeckel, David Hilbert, etc.

Esta anécdota resulta aquí esclarecedora por un doble motivo. Por un lado, nos sugiere que a finales del siglo XIX, precisamente cuando se constatan avances serios en la localización de las funciones cerebrales (gracias a los trabajos de Paul Pierre Broca y Carl Wernicke), se asientan las bases de la neurofisiología (con el propio Emil Du Bois-Reymond a la cabeza, sin olvidar a Johannes Müller o a Ernst von Brücke) y se abre un nuevo horizonte de comprensión del funcionamiento neuronal y del sistema nervioso (con los estudios de Wilhelm His, Auguste Forel, Camillo Golgi o Santiago Ramón y Cajal), aflora también, aunque pueda parecer paradójico, la sospecha de que seguir sin ningún atisbo de crítica la senda estrictamente fisicalista (en cualquiera de sus más diferentes variantes) a fin de esclarecer de una vez por todas la naturaleza de la consciencia, conduce irremediablemente a un sombrío callejón sin salida. A lo anterior cabe añadir que, desde entonces, este debate no se encuentra ni mucho menos zanjado y que se muestra, ya a partir de la segunda década del siglo XX, con renovada fuerza.

El filósofo Kim (2000), a este respecto, no duda en señalar a Feigl (1958) y Smart (1959) como los principales promotores modernos de la teoría de la identidad mente-cuerpo que, si bien ha tenido una preeminencia efímera, ha contribuido decisivamente a establecer y apuntalar los parámetros reflexivos desde los que, en la actualidad, se encauzan los debates más sobresalientes en el campo de la filosofía de la mente. Aun así, el dominio de lo mental sigue siendo un desafío para el conocimiento científico en lo que se refiere a su descripción física o fisiológica. Kim (2005) trae a colación de manera premeditada la dimensión fenoménica, subjetiva o cualitativa de los estados mentales (es decir, los «qualia» que algunos representantes de la filosofía materialista como Daniel Dennett, Michael Tye o Paul Churchland, plantean suprimir por superfluos) para defender con firmeza la inviabilidad epistemológica de reducirlos a parámetros estrictamente fisicalistas. Por mucho que se nos haga ver el cerebro como el trasfondo material desde el que «sobreviene», se «realiza» o «emerge» lo mental, no se ha logrado conjurar en esencia, y a pesar de todas rimbombantes promesas que se anuncian en selectos encuentros científicos, lo que de esquivo contiene. Y este escollo, no es posible negarlo, impele en gran medida a considerar la mente como algo anómalo, como una realidad perturbadora e incómoda.

Pues bien, es en este contexto particular donde reemerge con inusitada pujanza el discurso, un tanto nebuloso, del transhumanismo y su ofuscado interés por la actividad de la mente, hasta el punto de que, haciendo nuestro lo dicho por Tallis (2008), nos atreveríamos a afirmar, sin cometer una exageración impropia, que se encuentra penetrado de una

abrumadora neuromanía³. Desde ese punto de vista, los postulados del transhumanismo podrían ser interpretados como un flamante y llamativo abordaje del irresuelto problema sobre el nexa ontoepistemológico entre mente y cerebro, que pone todas las certidumbres en los avances de la bioingeniería y en el barniz evolutivo con el que se maquillan las trayectorias tecnológicas de nuestra contemporaneidad, especialmente aquellas que se atribuyen a la inteligencia artificial (IA). Esto no es algo novedoso. Existe una larga tradición en la cultura occidental, desde el pretérito acervo sapiencial relacionado con la creación de homúnculos y entidades golémicas⁴, pasando por la fascinación romántica por los autómatas, y terminando en las turbadoras especulaciones de la ciencia ficción moderna⁵, que ha concentrado algunos de sus temores más «siniestros» (de acuerdo con la conceptualización freudiana)⁶ o, por el contrario, sus ambiciones más intrigantes en el supuesto advenimiento de la emancipación maquinal.

La clave aquí tiene que ver con la íntima presunción de que el modo de conocimiento que atesora la empresa tecnocientífica puede solventar de un plumazo el viejo problema de la sustancialidad de la mente, de su sede o localización genuina. El avance de lo tecnológico ofrece a los ojos de algunos una retórica escatológica efectiva centrada en la sustitución de la organización material de la realidad (de redes neuronales se pasa a circuitos integrados), sin renunciar a los ya manidos argumentos del fisicalismo tradicional. Esta artimaña, más cercana a las refinadas estratagemas de la propaganda que a lo estrictamente científico, permite despachar el complejo asunto de la consciencia mediante su entrada en la horma de la teoría computacional. Pero no solo eso. El poder de la tecnología informática presenta un efecto disolvente de la ontología tradicional mediante la reconversión de los sistemas físicos en estructuras cuya interacción (tanto a nivel interno como con el entorno) se basan en el procesamiento de información. Una vez hecha esta transposición de sentido, aplicado a todos los órdenes de la realidad (incluidos los fenómenos vivos), se abren las puertas de par en par para una heurística omniabarcante, una teoría del todo con base en un orden inédito, opuesto al paradigma mecanicista, caracterizado por el «pancomputacionalismo» y el «paninformacionismo». El impulso de esta semántica de la información, bajo los parámetros de un naturalismo evolutivo, permite todo tipo de especulaciones, como la de la física «digital» —defendida, entre otros, por Edward Fredkin, Tommaso Toffoli, Stephen Wolfram, John Wheeler⁷— que explora la posibilidad de que el universo pueda ser el resultado de un programa computacional determinista o probabilístico (y, por tanto, que funcione como tal) o, incluso más allá, que sea una computadora que procesa un programa instalado desde el

³ De acuerdo con Tallis (2008), el dogma de la «neuromanía», cada vez más extendido en aquellos foros públicos y académicos en los que se defiende una narrativa naturalista, parte de la idea engañosa de que el ser humano, como ser vivo consciente, se reduce a los patrones de actividad cerebral.

⁴ Para seguir el rastro de esta legendaria figura es obligado remitirse a los textos de Moshe Idel, André Neher o Gershom Scholem. La interpretación literaria moderna del mito la encontramos mejor representada, sin duda, en la enigmática obra homónima de Meyrink (2010).

⁵ Véase al respecto, Sierra (2021).

⁶ Hacemos mención, claro está, al concepto de *Unheimlich* (contenido en el famoso texto *Das Unheimliche* de 1919), que el padre del psicoanálisis lo aplica para referirse a algo espeluznante y angustioso, próximo a la vivencia personal (aunque oculto), y que finalmente sale a la luz.

⁷ Para ser precisos, la hipótesis de que el universo es un ordenador digital fue propuesta inicialmente por el ingeniero alemán Zuse (1969).

principio de los tiempos (sin que, por el momento, que tengamos noticias del «gran programador» inicial).

El caso es que ni la irrupción de este orden ontológico digital, en la que descansa la teoría de la simulación, ni los espectaculares avances en el campo de la neurotecnología de los que estamos siendo testigos hoy en día, han contribuido de manera trascendental a desatar el nudo gordiano que atenaza tradicionalmente el escurridizo problema de la consciencia y de la relación mente-cerebro. Frente a ello, el transhumanismo ha tomado partido por una visión cientificista-tecnicista, al creer, a pies juntillas, en la omnipotencia de la racionalidad operativa que suele acompañar a la metafísica tecnológica contemporánea. Hay aquí un trasfondo de optimismo ingenuo que tiene que ver con el aura hipnótica y cuasi-mágica atribuida al avance de la robótica (siguiendo la tercera ley de Arthur C. Clarke)⁸, con la cualidad revolucionaria que se atribuye al fenómeno tecnológico y que, en última instancia, oculta una senda ya trillada (*Nihil novum sub sole*). De esta forma, en tanto que el transhumanismo no se sale del redil construido por el fisicalismo neurobiologicista (eso sí, aderezado de una «elocuente» antropología tecnomaterialista), nos hallamos simplemente ante una versión remozada que redundante en antiguos atolladeros filosóficos. Y esta escenificación «gatopardiana», en definitiva, no consigue disipar la sensación de que la evocación exaltada del *homo technologicus* nos sitúa en el mismo callejón sin salida anunciado por Emil Du Bois-Reymond hace más de ciento cincuenta años.

CUANDO LOS ÁRBOLES NO DEJAN DISTINGUIR EL BOSQUE. LOS ESPEJISMOS MITOPOIÉTICOS DEL TRANSHUMANISMO

El ascenso y el crédito creciente del transhumanismo en aquellos cenáculos de reflexión que trazan meticulosamente, con compás, escuadra y cartabón (ahora, en cambio, estaríamos hablando de modelos informáticos proyectivos), el futuro derrotero de nuestras sociedades constituye un indicador verdaderamente esclarecedor y ajustado sobre el modo en que se desea presentar la ciencia en el espacio de lo público. En este caso, encontramos que la certera representación o puesta en escena de su práctica, esto es, la estética tecnocientífica, alcanza aquí un poder mitogénico inaudito y se convierte, por sí mismo, en un mecanismo de legitimación central. Es más, el desmedido poder de sugestión, la potencia persuasiva del transhumanismo (transformado ya en un vaporoso imaginario de evocaciones utópicas) se ve acrecentado exponencialmente conforme se extiende el efecto de oscurecimiento que lleva a cabo sobre la propia actividad científica. Asistimos a una situación paradójica en la que la ciencia acaba siendo víctima de su propia aura proyectiva, es decir, de la envoltura de especulación visionaria con la que se arroja, cuyos tropos discursivos más relevantes se hallan condicionados por esta corriente de euforia desmedida en pos de lo tecnológico. En todo ello, aunque no es el único factor en juego, algo tiene que ver la función que ejercen determinadas estrategias dominantes de la comunicación científica en el disciplinamiento de la percepción y en el modo de adquisición de información por parte de las audiencias

⁸ Concretamente, es la siguiente: cualquier tecnología lo suficientemente avanzada es indistinguible de la magia (Clarke, 1973).

públicas a través de motivaciones que podrían considerarse «tendenciosas» o «sesgadas». Esta circunstancia no ha pasado en modo alguno desapercibida para la perspectiva transhumanista, cuyo alcance promocional es resultado, en el mejor de los casos, de exposiciones grandilocuentes que van a contrapelo de una información veraz y terminan convirtiéndose en mera propaganda de «hitos» científicos ligados a empresas tecnocientíficas privadas. Desde este punto de vista, la retórica transhumanista encuentra un grato acomodo en las lógicas de espectacularización científica, en la exhibición hiperbólica propiciada por el *marketing* del conocimiento, en los intereses que subyacen bajo la retórica agnotológica y, cómo no, en las férreas sujeciones de una economía de la atención.

Pero podemos incluso ir más lejos. El transhumanismo posee también la fascinante capacidad de escorar ideológicamente la presentación de la actividad científica, de tal manera que se convierta en la viga maestra sobre la que se construye un nuevo mundo que orbita obsesivamente en torno a lo que Morozov (2015) ha tenido el acierto de identificar como «solucionismo tecnológico»⁹. Este es un asunto clave que, si bien ha sido tratado con asiduidad, no ha sido todavía calibrado en su dimensión más compleja. Entre otras razones, porque es necesario hacer hincapié en el hecho de que nos hallamos ante un universo de conocimientos cuyos rasgos epistemológicos y axiológicos son muy diferentes si los comparamos con el conocimiento científico en un sentido, digamos, clásico. Tal y como nos lo hace saber el filósofo alemán Nordmann (2012), el conocimiento tecnocientífico, al que se adhiere abiertamente el transhumanismo, sitúa el foco de interés en la viabilidad de capacidades técnicas adquiridas y no en la confirmación o refutación de hipótesis sobre el mundo que nos rodea. Dicho con otras palabras, la validez de la tecnociencia se basa en la capacidad de extraer hábitos de acción de la cultura instrumental existente para pasar a operar, bajo protocolos previamente fijados, en ciertas propiedades específicas de la realidad y establecer sobre ellas procedimientos de control. No hace falta ir mucho más lejos en nuestro sobrio análisis para percatarnos de que, en el modelo tecnocientífico, se ha consumado un cambio trascendental respecto al tratamiento y la propia comprensión del objeto de investigación. De este hecho concreto se pueden extraer consecuencias que exceden el dominio de la filosofía de la ciencia y que se asientan, más bien, en el terreno sociopolítico o económico, dado que, en términos generales, acarrea un entendimiento distintivo y sin precedentes del mundo.

Pues bien, uno está obligado a familiarizarse en detalle con este horizonte epistémico emergente si desea interpretar adecuadamente la ofuscada fijación neurocéntrica de los grandes representantes del transhumanismo. El cerebro se transforma aquí en un adversario formidable que resulta preciso descifrar bajo parámetros de dominio instrumental para hacer realidad esa especie de «antropotécnica» de la que habla Sloterdijk (2000, 2012)¹⁰ y, en definitiva, consumir las viejas y no del todo satisfechas aspiraciones emancipatorias frente a toda constricción corporal o escollo biológico. La clave está precisamente en preparar las

⁹ Se trata, solo en apariencia, de una novedosa ideología global que defiende que todos los problemas de la humanidad pueden solucionarse con algún tipo de sistema o artefacto tecnológico.

¹⁰ El pensamiento de Sloterdijk sobre el fenómeno técnico parte, en gran medida, de sondear el soterrado impacto de esta (ya sea como «homeotecnología» o como «alotecnología») sobre el proyecto antropotécnico occidental (como «mejora del mundo» [*Weltverbesserung*] y como «mejora de uno mismo» [*Selbstverbesserung*]).

condiciones concretas para propiciar un horizonte venidero centrado en la «evolución dirigida»¹¹. Con ello, se sugiere sin atisbo alguno de modestia que la tecnociencia puede aspirar a asumir el papel protagónico en el proceso de transformación de la materia viviente. Y el hombre, en resumidas cuentas, se convierte en agente causal de su propio destino biológico que, de este modo, se abre a un sinfín de posibilidades en cuanto a la capacidad de llevar a buen término todo tipo de síntesis artificiales. Este discurso se ha visto remarcado básicamente en ciertas líneas de avance que se están gestando en diversos campos de la neuro-tecno-ciencia de vanguardia como la neuroingeniería o la neurobiotecnología, ya que los llamativos logros provenientes de los sistemas de convergencia artificial mente-máquina (BCI), de la neuromodulación artificial y, en fin, de las tecnologías de intervención directa sobre localizaciones y conjuntos de conexiones sinápticas o circuitos neuronales del cerebro (por ejemplo, las técnicas de ultrasonido enfocado de baja intensidad, LIFUP, por sus siglas en inglés) han tenido como consecuencia avivar la controversia filosófico-epistemológica sobre una eventual e inédita transformación de la autognosis corporal y de la percepción de la realidad.

En este sentido, el transhumanismo, es preciso reconocerlo, ha tenido la enorme virtud de detectar con agudeza el alcance y las potencialidades del modelo aplicado de conocimiento que deposita su absoluta confianza en la ingenierización o control operativo del cerebro (y que, a un corto plazo de tiempo, podría conducirnos hacia un escenario factible de intervención de su fisiología a través de la implantología neural, de la reprogramación de circuitos neuronales mediante optogenética¹², del desarrollo de dispositivos neuromórficos¹³ o del cultivo controlado en laboratorio de cerebroides¹⁴). Pero aquí hay un matiz que no puede pasar desapercibido. En este contexto, lograr cartografiar la totalidad del territorio cerebral, deshilar la compleja maraña dinámica del conectoma neuronal humano¹⁵ y, en suma, desvelar los misterios más ocultos de su funcionalidad interna son progresos previsibles que, paradójicamente, sirven para remarcar su depreciación y su inevitable caducidad. Insistamos en ello porque es crucial. No es el cerebro, sino la neuro-tecnología la que se abre paso para el cumplimiento de la escatología transhumanista. Y para ello es necesario una rearticulación o recomposición tecnológica de la función cognitiva que, como antesala

¹¹ Con este concepto se hace referencia a un método utilizado especialmente en el campo de la ingeniería de las proteínas y en la biotecnología basado en la imitación del mecanismo de la selección natural con el propósito de orientar o conducir la evolución de proteínas o ácidos nucleicos hacia objetivos definidos.

¹² Este campo tecnocientífico emergente es el resultado de combinar métodos genéticos y ópticos (por lo general, destellos de luz a diferente frecuencia) con el objeto de controlar, activar o modificar, mediante la introducción de genes exógenos que codifican proteínas fotosensibles, comportamientos específicos en ciertos grupos de neuronas.

¹³ Digamos que la ingeniería neuromórfica o cómputo neuromórfico es un concepto desarrollado, a finales de la década de los 80 del siglo pasado, por el informático Carver Mead que parte del estudio de la morfología, estructura y adaptabilidad de las neuronas con el objeto de simular o incluso diseñar sistemas neuronales artificiales.

¹⁴ Los organoides son cultivos tridimensionales originados de células madre pluripotentes que poseen estructuras y funcionamientos similares, en una escala reducida, a los propios órganos de los que se ha extraído mediante biopsia (colon, hígado, páncreas, estómago, riñones, retina, tracto gastrointestinal, corazón y cerebro).

¹⁵ A grandes rasgos, un conectoma es un mapa de las conexiones existentes entre las neuronas del sistema nervioso de un organismo. La producción y el estudio de los conectomas, entendidas como redes dinámicas de conexión, se conoce como conectómica.

de una fase ulterior, conduzca a la definitiva erradicación del propio cerebro y su sustitución por sofisticados ingenios de inteligencia artificial.

Se repite una vez más, esta vez actualizada bajo la égida de una especie de utopismo religioso tecnófilo, la idea (elevada a ensoñación metafísica) del cuerpo como residuo, prisión, cloaca o incluso como sepulcro, cuyas raíces más profundas se hallaban disimuladas en el cosmos noético platónico y encuentran en el cristianismo una acogida sin precedentes cuando se impone el mandato paulino de renunciar a la condición de *ἄνηρ σαρκός* (hombre carnal) en pos de experimentar una metanoia purificadora que provoque el surgimiento del *ἄνηρ πνευματικός* (hombre espiritual)¹⁶. Este desprecio de lo carnal, que arriba hasta la modernidad a través de diversas ramificaciones, adquiere ahora algunos matices realmente desconcertantes con la escisión tecnocientífica de la condición humana en dos nuevas realidades ontológicas mutuamente complementarias, aspecto que ha sido observado con tremenda lucidez por Günter Anders. El «homo creador», capaz de transformar todas las cosas de manera sustancial, incluido él mismo, se constituye, al mismo tiempo, en «homo materia», es decir, en la «materia prima de sus propias producciones tecnológicas» (Anders, 2011). El ser humano, así, viene a convertirse en el principal objeto de transformación y dirección tecnológica.

En lo que se refiere a esta reificación incontenible de la naturaleza humana, expresada en el sistema nervioso central, es sabido que los hitos de intervención y estimulación tecnomédica en la topología fisiológica cerebral, aunque se adorne con todo tipo de oropeles mercadotécnicos, no son ni mucho menos novedosos y se retrotraen unas décadas en el tiempo. De hecho, podemos remontarnos a los años 50 del siglo XX para localizar los primeros ejemplos de neuromodulación artificial en laboratorio. Ciertamente, el potencial práctico y el alcance insospechado que encierra el mecanismo artificial de transformación del mundo a través del cerebro se expone de forma inquietante en los experimentos realizados por Olds y Milner (1954) sobre el septo cerebral (*septum pellucidum* o septo lateral transparente) de ratas, una zona particular que, al ser estimulada eléctricamente, segrega una importante descarga de dopamina (provocando intensas sensaciones de placer y excitación sexual). Por supuesto, este tipo de experimentos no tardaron en llevarse a cabo en pacientes humanos. Por ejemplo, hay que destacar en este punto el famoso trabajo en el campo de la psiquiatría biológica del neurocirujano Heath (Bishop et al., 1963), que partía del implante de electrodos y de la aplicación de descargas eléctricas para erradicar la homosexualidad (es muy conocido, al respecto, el famoso caso del paciente «B-19»), o la invención del chip cerebral y de otros dispositivos neuroestimuladores (como el «*stimoceiver*») por parte de Delgado (1969). Todo ello anticipa una línea de investigación en el campo de la tecnociencia de vanguardia que acoge con entusiasmo los avances en la neuromodulación para corregir eficazmente los trastornos cerebrales de la epilepsia o la depresión, pero también para aumentar significativamente la capacidad cognitiva del ser humano («aumento artificial») a través de una estrecha exposición a reguladores electrónicos (haciendo uso de la mera estimulación,

¹⁶ No han sido escasos los filósofos (como Félix Duque, Iñaki Arzo, Andoni Alonso, Josep María Esquirol y tantos otros), que han descubierto la íntima conexión entre cierta metafísica etérea o, incluso, cierta sensibilidad sacramental y el modelo antropológico que se desprende de la ingenierización radical de la vida.

mediante la implantación de chips integrados o incluso aplicando, como ya se ha mencionado anteriormente, técnicas de ultrasonidos focalizados de baja intensidad).

Con todo, no nos llamemos a engaño. Las promesas y pronósticos que se lanzan alegremente a los cuatro vientos por el transhumanismo son, si nos paramos a pensar en ello, ostentosos brindis al sol. Esto no puede extrañar a nadie a estas alturas, ya que las suplementaciones tecnológicas que se auspician enfáticamente desde el transhumanismo insinúan en realidad una aspiración por tomar el control operativo de lo que es, y aquí hacemos un ejercicio de contención expresiva, muy difícil de controlar. Y eso supone inevitablemente incorporar aproximaciones harto simplificadas de la fisiología cerebral y de la propia consciencia que son fruto de un cientificismo galopante y de un neurocentrismo autorreferencial. La tesis de que la naturaleza de la consciencia o el pensamiento puede ser explicada por completo con el refinamiento exploratorio de la neurociencia no deja de asombrar, por mucho que se insista en ello, cuando sabemos que componentes fundamentales de la mente y del «yo» se hallan «externalizados» y dependen de condiciones sociales, culturales y vivenciales concretas. Más aún, esta lamentable confusión suele trasladarse también a la investigación estrictamente «objetiva» del cerebro humano cuando se nos hacen pasar ciertas abstracciones anatómicas obtenidas mediante imagenología cerebral por reflejos exactos del funcionamiento neuronal. No es comparable el registro digital de unas cuantas decenas de miles de trayectorias sinápticas neuronales —como el conseguido por científicos de Google con el hemisferio de la mosca de la fruta, (*Drosophila melanogaster*)—¹⁷, con la inmensidad dinámica del cerebro humano, ya que despliega cotidianamente una cantidad de combinaciones mayor que el número calculado de átomos existente en el universo¹⁸. Volveremos más tarde sobre este asunto.

Ahora bien, lo que no se puede perder de vista aquí es que, teniendo en cuenta que la intervención tecnomédica que se realiza sobre la actividad fisiológica de las estructuras neuronales puede provocar efectivamente una profunda alteración de la subjetividad humana, estamos asistiendo a un avance trascendental que puede ser susceptible de ser aplicado a las políticas estratégicas de la ingeniería social contemporánea. Así las cosas, corremos el peligro real de pasar de un conductismo persuasivo a nivel psicosocial (como lo propugnaban Gustave Le Bon, Wilfred Trotter, Leon Festinger o Edward L. Bernays) a un modelado tecnomédico del comportamiento colectivo. Esta posibilidad, que está siendo confirmada en parte cuando la OTAN (Organización del Tratado del Atlántico Norte) nos habla ya de la incorporación de la llamada «guerra cognitiva» en sus estrategias de contienda híbrida y del cerebro como «el nuevo campo de batalla del siglo XXI», nos sitúa

¹⁷ Científicos de Google y del Janelia Research Campus de Virginia (EE. UU) han publicado el mayor mapa de alta resolución de la conectividad cerebral en cualquier animal, compartiendo un modelo 3D que traza 20 millones de sinapsis que conectan unas 25 000 neuronas del hemisferio de una mosca de la fruta.

¹⁸ El cerebro humano contiene alrededor de 86 000 millones de neuronas. Cada neurona posee decenas de miles de conexiones neuronales y se estima que se pueden llevar a cabo 100 billones de sinapsis neuronales. La cantidad de almacenamiento necesaria para albergar esta información cada segundo para una persona sería de más de 100 000 terabytes o 100 petabytes. Pero este cálculo no contempla los cambios en la conectividad y posicionamiento de estas sinapsis a lo largo del tiempo. Contar los cambios de estas conexiones, justo después de una noche de sueño o de una clase de matemáticas (tal y como ha sido señalado por numerosos expertos), equivaldría a una cifra enorme que se aproxima a los 10^{80} átomos del universo.

dentro de una tradición distópica alimentada por destacadas obras literarias que han forjado una influyente imaginaria vinculada a la tecnociencia contemporánea como *La isla del doctor Moreau* (Wells, 1896), *Nosotros* (Zamiatin, 1924) o *Un mundo feliz* (Huxley, 1932). En estos textos, la intervención tecnomédica en el cerebro desempeña un papel destacado en la conformación de un modelo paradigmático del ser humano, caracterizado por la entrega ilimitada al escapismo ficcional, el hedonismo ensimismado y el consumo compulsivo de sustitutos sedantes frente a la experiencia genuina.

JUEGOS DE ENMASCARAMIENTO. LA ANALOGÍA CEREBRO-COMPUTADOR

El transhumanismo constituye uno de los legatarios más descollantes del relato mítico en torno al fértil universo de sentido atribuido a la metáfora maquina que autores como Dijkstra (1969) o Mumford (2016) han situado en el centro de la cosmovisión occidental. Ciertamente es que la interpretación de lo artificial subyacente al *corpus* alegórico transhumanista y remite a una racionalidad pragmático-instrumental (derivada del avance de la bioingeniería y de la inteligencia artificial) algo alejada de la filosofía mecanicista renacentista (Brunelleschi, Fioravanti, Taccola, Fontana, Kyeser, Da Vinci, etc.) o, incluso, de la propia cultura mecánica que irrumpe con todo su esplendor en el pensamiento occidental allá por los siglos XVII y XVIII (Descartes, Bacon, Hooke, La Mettrie, etc.). No obstante, en su trasfondo más reservado, en el estrato más profundo, persisten todavía, a pesar de la «distancia en el tiempo» (si bien camuflados con otras desafiantes y prodigiosas máscaras), todo un elenco de elementos metafóricos, imaginarios y tradiciones especulativas que vienen a forjar la alegoría antropopéyica contemporánea (Ball, 2012). De esta forma, la imperecedera analogía que enlaza la vida con el mecanismo se entremezcla, a su vez, con otro tipo de variantes asociativas de mayor refinamiento técnico, tan afines al pensamiento transhumanista, como la que equipara el cerebro humano (y con él, la mente) con un computador. Con base en esta tesis, construida fundamentalmente sobre supuestos y figuraciones asociadas a la inteligencia artificial, se han dado múltiples posicionamientos acerca de la naturaleza de la consciencia que oscilan, *grosso modo*, desde un dualismo *sui generis* tendente al funcionalismo informático (o lo que es lo mismo, la mente como un *software* que se ejecuta en un cerebro-computadora) hasta una especie de neuromonismo radical que sostiene la identidad del fundamento material y la consciencia.

He aquí el *quid* del asunto. Porque, aunque pueda parecer sorprendente, el hecho concreto es que muchos de los protagonistas que, de una u otra forma, han terciado en la fundación y desarrollo de la IA se han dejado arrastrar, en el fondo, por la maleabilidad y ductilidad de los tropos narrativos asociados a la teoría computacional de la mente para crear novedosos y encomiásticos mitologemas (haciendo nuestro el concepto de Kerényi)¹⁹. Ahora bien, aquí la carga de ensoñación quimérica no radica en la posibilidad o no de que la IA tenga lugar en alguna etapa determinada del progreso humano, sino precisamente en el carácter de

¹⁹ Aludimos con este concepto a la constante transformación del material mítico mediante la provisión de contenidos de una cultura concreta. Véase al respecto, Jung y Kerényi (2004).

inevitabilidad que algunos de sus partidarios más exaltados atribuyen a su advenimiento. Con estos mimbres, se demuestra bien a las claras la existencia de una óptica colmada de cierta inverosimilitud en los espacios internos en donde se gestan las estrategias de la praxis tecnocientífica, en la medida en que se persigue de manera recurrente y, a nuestro entender desacertada, la emulación escrupulosa de características particulares de la inteligencia con el propósito preciso de alcanzar la denominada inteligencia artificial general (IAG). Pero no solo eso. Sería completamente disparatado dar la espalda, en esta época tan excepcional de apoteosis computacional, al influjo de un imaginario adicional que se arraiga, cada vez de modo más tenaz, en el substrato cultural de nuestras sociedades contemporáneas y que tiene que ver con el orden anunciado de una fase evolutiva concluyente del desarrollo tecnológico (y aquí entran autores paradigmáticos del transhumanismo como Ray Kurzweil, David Andrew Pearce o Nick Bostrom o Elon Musk) que, además de periclitarse aquellas propuestas, en parte más realistas o, si se quiere, más bajadas a tierra, relacionadas con la IA, abren un horizonte de prospectiva futurista atravesado por una dialéctica que confronta una perspectiva tecnooptimista con otra más sombría y distópica.

A decir verdad, este imperativo tecnológico avizorado por el transhumanismo soslaya con gran habilidad el hecho de que las posibilidades de desarrollo de la IA se encuentran limitadas por el propio nivel de conocimientos que se atesoran en la actualidad sobre la naturaleza de la inteligencia humana. Por mucho que se haya insistido sobre ello, las líneas de trabajo en torno a la IA que se están llevando a cabo desde los años cincuenta se obstinan en el desacierto de basarse en un modelo altamente simplificado de la inteligencia humana. Se trata de una senda tortuosa que no lleva a puerto seguro porque, pese a los múltiples intentos operativos de imitación o emulación de pensamiento, hay una cosa que parece clara: la inteligencia humana y la artificial son radicalmente desiguales (Dreyfus, 1992; Dreyfus y Dreyfus, 1986; Nöe, 2009; Larson, 2021). No obstante, las preocupaciones y áreas de atención prioritarias que irrumpen con la aparición de la IA (cuyo germen incipiente data de la famosa Conferencia de Dartmouth, de 1956, donde se reúnen algunos de los grandes popes de esta ciencia en ciernes como John McCarthy, Marvin L. Minsky, Nathaniel Rochester y Claude E. Shannon) se han tratado de abordar a través de una sistemática identificación de problemas particulares que se sacan fuera del contexto real en el que se produce la inteligencia humana y a través de la reducción de esta a mera potencia computacional, lo que explica, en última instancia, el triunfo actual de una enfervorizada mentalidad «dataista» (Han, 2014). Pero dejemos claro que esta espinosa problemática no es de ahora. Debemos trasladarnos a las raíces iniciales del proyecto de realización de máquinas inteligentes, impulsado por el matemático británico Turing (1950) y, posteriormente, por el lógico y filósofo austriaco Gödel (1931) —con su archiconocido principio de incompletitud de los sistemas formales— para atestiguar el surgimiento de un equívoco trascendental que iba a prosperar en las siguientes décadas. En la firme pretensión de reproducir maquinalmente las facultades intelectivas del pensamiento humano, se ha considerado adecuado llevar a cabo extrapolaciones basadas en reglas, operaciones estadísticas o estrategias de resolución de problemas. Huelga decir que la clave de bóveda aquí es de carácter cualitativo y no cuantitativo. Pese a la eficacia retórica que acompaña a ciertas campañas de comunicación científica, el avance en el perfeccionamiento de los sesgos funcionales aplicados, en los métodos modernos de aprendizaje automático o en el aprendizaje profundo resulta, en este caso, absolutamente secundario. Esto no conduce al

logro de una inteligencia del nivel del ser humano, y si lo hace, insistimos en ello, a un callejón sin salida para la inteligencia artificial.

La razón de esta afirmación tan contundente resulta hasta cierto punto comprensible. No en vano, la IA trabaja con una concepción de inteligencia estrictamente universal, proyectada en un marco de pura abstracción, sin percatarse de que la inteligencia humana descansa siempre y necesariamente en un contexto sociocultural concreto. De ahí que las estrategias y los métodos de programación de la IA, aspecto fundamental relativo a la manera en que esta rama de la ciencia computacional interacciona con la realidad, posean un alcance reducido en la cuestión del alcance cognitivo de la deducción (o sea, en la generación de conocimiento) y de la causalidad (esto es, en la ordenación y dotación de criterios de relevancia para nuestros conocimientos sobre el mundo). Lo que se está dirimiendo en realidad es, en resumidas cuentas, si es factible trascender el mero cálculo e incorporar a los dispositivos artificiales aquellas conjeturas particulares que se nutren del sentido común cultivado y percibido por el hombre en su vida cotidiana. Y esto es algo que tiene su importancia porque, hoy en día, el proceder de la IA se atiene al radio de acción de enfoques deductivos de la inferencia o, ya a partir de los años noventa, a un patrón de inferencia deductiva que en ningún caso escapa de los marcos lógicos previamente fijados. El entorno en el que se desenvuelve la IA es, en definitiva, cerrado y sometido a un control mayúsculo y completamente previsible, algo opuesto a la experiencia humana de la realidad que, acláremoslo por si hubiera alguna duda, es abierta, provisional y en continua transformación. En ese sentido, el procesamiento de datos y los análisis que de ellos se derivan a partir de ciertas operaciones de simulación de comportamiento (función realizada desde el aprendizaje profundo) o de la inducción automatizada resultan, cuanto menos, escasos y rudimentarios a la hora de abarcar y manejar la ingente cantidad de variables ocultas que son contempladas de forma natural por el ser humano durante su vida en los procesos de «inferencia subyacente» o de «inferencia abductiva» —por utilizar la terminología de Peirce (1929)—. A diferencia de lo que ocurre con normalidad en el ser humano (cuyo conocimiento se abastece de aquel caldo de cultivo que comúnmente se denomina «sentido común»), la IA no ha dado resultados exitosos, más bien lo contrario, a la hora de integrar esta cualidad intelectual, esta especie de instinto, de selección enigmática, de salto cualitativo, de conjetura explicativa que se nutre de las condiciones específicas del contexto y de la infinidad de posibilidades que este encierra (Levesque, 2014). He aquí una divergencia, por el momento, insalvable con el pensamiento humano y que retrata una realidad, si se quiere, más modesta, ya que, ante la ausencia del trasfondo contextual, los sistemas de IA no resultan en modo alguno eficaces para establecer inferencias, contextualizar el conocimiento y sacar conclusiones relevantes.

Hay otros aspectos sobre la IA vinculados a lo dicho hasta ahora y que, más que asentarse en un terreno sólido y estable, siguen abundando en un craso error teórico-conceptual y en ciertos espejismos fantasiosos. Señalemos dos ejemplos significativos. En primer lugar, está la aspiración, celebrada como algo seguro por los apologetas transhumanistas más entusiastas, para lograr una superinteligencia artificial (Szocik et al., 2020). Se supone además que, al introducir la retroalimentación en la mejora maquina —es decir, la automejora, concepto que, por cierto, fue rechazado por el propio Neumann (1966) décadas atrás— se origina una curva exponencial de inteligencia en las máquinas que desemboca, en su tramo final, en un

estadio irreversible de «ultra-inteligencia». No hay que ir muy lejos para imaginar que, con esto, se están poniendo las bases para anticipar, con visos de realidad y en un horizonte temporal relativamente temprano, la superación del pensamiento humano —aspecto que es retomado en la actualidad por Bostrom (2017)—. Sin embargo, aquí subsiste, de nuevo, una confusión elemental. Por mucho que se potencie el *hardware* en la carga y ejecución de aplicaciones (cumpliéndose así la Ley de Moore²⁰), esto no se asemeja, ni de lejos, a una inteligencia consciente. Las capacidades intelectivas del ser humano escapan a todo intento de mecanización reductiva. En segundo lugar, está la idea de la singularidad tecnológica que, en un principio, el matemático Vinge (1993) traslada a la computación y, específicamente, a la inteligencia artificial. Como es bien sabido, será finalmente el prospectivista Kurzweil (2005) quien da un empuje renovado a este concepto con su «Ley de retornos acelerados», ampliando así el patrón de crecimiento exponencial en la complejidad de circuitos semiconductores y pregonando a los cuatro vientos el advenimiento próximo de la IAG, momento decisivo en el que supuestamente las máquinas y no las personas asumirán el control global como los seres más inteligentes del planeta. Sin embargo, cuando Kurzweil populariza este discurso ya hacía tiempo que había desaparecido en el campo de la IA la ilusión de alcanzar una superinteligencia (Larson, 2021).

Todo revela, en suma, que la tendencia del transhumanismo a simplificar la innata complejidad de la inteligencia humana, al subordinarla a la operatividad computacional contribuye, sin duda, a expandir una cosmovisión tecnocéntrica, centrada en una pretendida equiparación de la mente humana con la computadora. El escenario final arroja ciertas sombras amenazantes para el futuro de la ciencia y del reconocimiento del potencial intelectual de un ser humano que, bajo las derivas actuales de la computación basada en datos, puede terminar convirtiéndose en un engranaje minúsculo dentro de una maquinaria imponente. Atendiendo a todo lo dicho, no hay duda de que la IA, proyectada desde la óptica transhumanista (plagada de entelequias y fascinaciones quiméricas), se enfrenta, al igual que la parábola narcisista encarnada por Dorian Gray (el personaje creado por el escritor y poeta irlandés Oscar Wilde), ante un atrayente retrato distorsionado de sí mismo.

LA LICUEFACCIÓN DE LA REALIDAD. SOBRE ENTIDADES FANTASMÁTICAS DEAMBULANDO EN UN UNIVERSO DIGITAL

El transhumanismo, no solo pone en grave cuestionamiento el paradigma asociado a la localización anatómico-cerebral de la identidad humana y su consciencia (mente-consciencia), sino que recupera la antigua cuestión filosófica sobre la posible «inmaterialidad» del yo para barajar algunas alternativas especulativas que tienen que ver con la ubicación de la identidad y de la consciencia humana en la periferia del dualismo psicofísico y antropológico tradicional. Si la mente humana se caracteriza por una información emergente que surge de los patrones organizativos de una red neuronal, puede ser posible replicarla artificialmente en un medio

²⁰ Tesis predictiva que señala que cada dos años se duplica el número de transistores en un microprocesador.

físico diferente. Esta hipótesis rompe con el clásico dualismo cartesiano que Ryle (2005) va a ilustrar críticamente con el sugerente y original lema del «fantasma en la máquina»²¹.

La focalización en la actividad neuronal como condición suficiente (y no solamente necesaria) para la generación de las experiencias que se integran en la categoría de lo mental, abre un horizonte lleno de posibilidades en la producción artificial de la consciencia. Vayamos por partes. En primer lugar, en los últimos tiempos se empieza a sopesar con cierta circunspección, a medida en que avanza a marchas forzadas la ingeniería neural y la diferenciación dirigida en laboratorio de células madre, el experimento mental propuesto por Putnam en 1981 acerca de la posibilidad de mantener artificialmente la actividad de un cerebro separado del cuerpo y conservado en una cubeta (Putnam, 1988). Aquí la hipótesis cartesiana del genio maligno que interviene malévolamente en la percepción de la realidad del sujeto pensante (Descartes, 2009, 2011) se transforma en un computador controlado por algún ingenioso científico técnico (benévolo o maligno) que dota de experiencias ininterrumpidamente a un cerebro suspendido en una cubeta llena de líquido nutritivo. Dicho ejercicio mental pasa a mayores en el año 2019 cuando un grupo de investigadores, encabezado por el biólogo Alysson Muotri, anunció la creación de organoides cerebrales (cerebroides)²² que generaban ondas de actividad eléctrica coordinadas, con un patrón similar al observado en los electroencefalogramas de bebés prematuros.

Al margen de las cuestiones éticas (que constituyen harina de otro costal), todo esto saca a la palestra, una vez más, la cuestión relacionada con la adquisición de una consciencia inducida artificialmente y con el problema de la capacidad para distinguir experiencias propias del mundo frente a otras reguladas intencionalmente por mecanismos externos. A fuerza de ser reiterativo se ha de insistir, una vez más, en que dicho problema se plantea sobre premisas falsas, que descansan en un funcionalismo digital simplificador o en un reduccionismo fisicalista reductivo, al pretender que un cerebro aislado puede ser fuente única de la que emana una percepción coherente del mundo. El sentido, además de ser enderezado a través de condiciones materiales de carácter particular, se gesta mediante una red de significados que se construyen colectivamente, es decir, disposiciones que se articulan a partir de procesos sociohistóricos (García Selgas, 1994).

En segundo lugar, el ingenio visionario del transhumanismo da «otra vuelta de tuerca» en la alegoría computacional de la consciencia con una de las tentativas más pintorescas y «arriesgadas» para lograr una definitiva emancipación del yo de la evanescencia física del cerebro. Nos estamos refiriendo en este caso, claro está, a la tecnología de carga mental, también conocida como emulación de todo el cerebro (WBE). Se trata de un proceso teórico-futurista de escanear una estructura física del cerebro con la suficiente precisión como para crear una emulación del estado mental (incluida la memoria a largo plazo y el «yo») y transferirla o copiarla a un ordenador de forma digital. El ordenador ejecutaría, entonces, una simulación del procesamiento de la información del cerebro, de forma que respondería

²¹ Ryle (2005) presenta la teoría cartesiana como un mito filosófico que distorsiona nuestra comprensión de lo mental.

²² Los organoides cerebrales, algunas células gliales y sus neuronas se interconectan entre sí formando redes y generando o propagando impulsos nerviosos.

esencialmente de la misma manera que el cerebro original y experimentaría tener una mente consciente y sensible. La carga de la mente puede ejecutarse potencialmente por cualquiera de los dos métodos siguientes: el primero consistiría, a grandes rasgos, en copiar y cargar o copiar y borrar mediante la sustitución progresiva de las neuronas (lo que puede considerarse como una carga destructiva gradual), hasta que el cerebro orgánico original deje de existir y el segundo, en cambio, se apoyaría en un programa informático que emule el cerebro para tomar el control del cuerpo. La mente simulada podría estar dentro de una realidad virtual o un mundo simulado, con el apoyo de un modelo anatómico de simulación corporal en 3D. Otra posibilidad es que la mente simulada resida en un ordenador dentro de un robot (no necesariamente humanoide) o de un cuerpo biológico o cibernético. Sea como fuere, el propósito parece evidente. La inserción del estado mental emulado en un ordenador (ya sea mediante el escaneo y mapeo del cerebro orgánico original o mediante la sustitución gradual de las neuronas) pone en el horizonte de un futuro próximo la tesis de la ubicuidad cognitiva y, junto con ello, una probable modificación o incluso recreación artificial del ser humano —algo así como una «resurrección computacional» (Reichenbach, 1978) con la aspiración de aumentar las capacidades físicas, intelectuales y sensoriales, así como prolongar la duración de la existencia o controlar las emociones—.

Aunque es preciso interpretar en sus justos términos esta teoría como lo que es, es decir, un simple vaticinio presuntivo con una enorme carga ideológica, se ha de reconocer su utilidad para desmontar algunas de las creencias que acompañan el avance de la neurotecnología y que son tomadas como auténticos hechos consumados. Por una parte, ofrece un fértil terreno para entrar de lleno en la larga controversia entre el enfoque del localizacionismo anatómico²³ y la concepción «integral» centrada en el principio de equipotencialidad del tejido cerebral²⁴ que ha sido estimulada en años recientes con las técnicas de neuroimagenología (en concreto, tomografías axiales computerizadas, resonancias magnéticas, magnetoencefalografías, etc.). De aquí se ha deslizado subrepticamente la idea de que la tecnología ha alcanzado un grado tal de sofisticación que permite captar ya en tiempo real, no solo la actividad cerebral, sino también experiencias, motivaciones, intenciones, emociones o pensamientos, aspecto este que ha sido puesto reiteradamente en entredicho con argumentos sólidos sin que ello haya provocado una apasionada receptividad en los espacios de divulgación científica²⁵. Por otra parte, queda sin resolver la incógnita de si la copia digital de la consciencia escaneada o cargada tecnológicamente alberga la dimensión subjetiva que es esencial en la conformación de la identidad. Esto es imposible

²³ Tesis que defiende la existencia de funciones localizadas en zonas específicas de la fisiología cerebral.

²⁴ El principio de «equipotencialidad» postula que si ciertas partes del cerebro son dañadas, otras partes del cerebro podrían ocupar el rol de las partes dañadas, reconfigurar la estructura cerebral y reemplazar, a fin de cuentas, las funciones de las que se encargaban las estructuras cerebrales dañadas.

²⁵ Tallis (2016) resume esta postura crítica con los siguientes razonamientos: la resonancia magnética funcional no capta directamente la actividad neuronal, sino los flujos sanguíneos para la distribución de oxígeno; hay un retraso en la detección de los cambios de flujo sanguíneo respecto a la propia actividad neuronal; para detectar un cambio de flujo sanguíneo se deben activar muchos millones de neuronas, por lo que es posible que se distribuya oxígeno a más de una descarga neuronal; hay una infrarrepresentación en la imagen de pequeños grupos de neuronas que provocan cambios de flujo; no se activan partes específicas del cerebro ante un estímulo, sino también actividades adicionales; los cambios menores de actividad se pasan por alto; la identificación de la actividad neuronal adicional es resultado de un promedio que elimina las variaciones de respuesta ante estímulos sucesivos.

de verificar y remite, más bien, a otro experimento mental formulado, en este caso, por Chalmers (1996) —aunque podemos encontrar algunos antecedentes parciales de este planteamiento en autores tales como Saul Kripke, Thomas Nagel o Robert Kirk— y que aborda la probabilidad lógica de «zombis filosóficos», copias físicas de seres humanos carentes de consciencia.

Obviamente, no es posible detenerse aquí con minuciosidad en todas las implicaciones que cabe extraer de estas figuraciones hipotéticas. Sin embargo, conviene poner sobre la mesa un tercer tipo de abordaje que el transhumanismo cultiva con delectación a partir del tropo computacional. Se trata de la posibilidad real de modificar patrones adaptativos a través de la producción de una ontología simulada, en la medida en que se interviene en la dinámica de correspondencia entre los patrones de distribución espaciotemporal de la energía física (el estímulo proximal que es captado por los receptores sensoriales del ser humano) y la propia experiencia psicológica intransferible del individuo asociada a la interpretación de cada estímulo a la hora de dotarlo de sentido. No hay duda alguna de que esta premisa de partida, si se presta la atención adecuada, lleva a desvelar, como una certeza incontrovertible, una especie de ontología digital —la que identifica al universo con una computadora gigante, un autómatas celular (Zuse, 1969), una máquina de Turing universal (Hochreiter y Schmidhuber, 1997) o un computador cuántico (Lloyd, 2006)— que descansa en la hipótesis fuerte de la simulación o del universo holográfico²⁶. En resumidas cuentas, esta conjetura, que ha sido popularizada en nuestros días por el filósofo Bostrom (2003) a través de su famoso «trilema»²⁷ y Chalmers (2005) con su «hipótesis de la matriz»²⁸, parte de la idea de que la naturaleza del universo es discreta y, como tal, puede ser modelado numéricamente y, por tanto, computable como resultado de algoritmos.

Ciertamente existen múltiples críticas, de las más variopintas y con diferentes niveles de profundidad, a este argumento tan manido en la retórica transhumanista. Hay quienes han percibido, no sin razón, un trasfondo soteriológico, de corte más próximo al neoplatonismo o al neopitagorismo, en los niveles de jerarquía ontológica que se desprenden de la tesis de un universo simulado (Steinhart, 2010). Otros creen ver en todo esto una flagrante confusión entre la significación «predicativa» (es decir, la exploración del universo mediante herramientas computacionales) y «atributiva» (o sea, reconocer la naturaleza computacional de las leyes físicas) de lo digital a la hora de considerar la realidad, ya que esta «ontología

²⁶ Esta hipótesis está poniéndose a prueba a través de técnicas de cuadrícula de cromodinámica cuántica (un método por el que los superordenadores son capaces, por el momento, de simular minúsculas porciones del universo a escala de una mil millonésima parte de un metro, algo apenas mayor que un núcleo atómico).

²⁷ El trilema de Bostrom (2003) se basa en lo siguiente: la fracción de civilizaciones de nivel humano que alcancen un estado posthumano es muy cercana a cero; la fracción de civilizaciones posthumanas que están interesadas en ejecutar simulaciones de ancestros es muy cercana a cero; la fracción de todas las personas con nuestra clase de experiencias que están viviendo en una simulación es muy cercana a uno.

²⁸ Esta hipótesis consiste en lo siguiente: los procesos del mundo físico son fundamentalmente computacionales; nuestras mentes no forman parte del mundo físico, pero interaccionan con sus procesos; la realidad física fue creada por seres externos al espacio-tiempo físico.

artificial» choca frontalmente, en algunos casos, con algunos postulados descriptivos a nivel macroscópico de la ciencia contemporánea (Floridi, 2009)²⁹.

Sea como fuere, una cosa parece clara: la simulación es una atribución cualitativa de la realidad que se constata únicamente en relación con una referencia ontológica que no pertenece a aquel orden de lo real y con la que se mantiene algún contacto (Agatonović, 2021). Pero la cuestión no se detiene, ni mucho menos, en este punto. Hay implicaciones de alcance sociopolítico soterradas que arrojan una sombra inquietante sobre la autonomía y el libre albedrío de la condición humana. No solo se está planteando como posible producir una ontología simulada, de alterar artificialmente las condiciones externas del entorno mediante la incidencia en los estímulos proximales de los receptores sensoriales y, en consecuencia, de orientar deliberadamente los procesos de construcción cognitiva generados por el cerebro humano (algo parecido a los sistemas de interfaz cerebro-ordenador de la realidad virtual), sino de reencauzar a voluntad las trayectorias sinápticas que organizan nuestra percepción de la realidad. En suma, la tierra prometida profetizada por la metafísica transhumanista quebranta la solidez de cualquier referente perceptivo y nos sumerge en un sinuoso e impredecible «mundo de visiones fantomáticas» (de acuerdo con el concepto imaginado por Stanisław Lem), donde la retórica fabulosa de lo onírico, aquella que soliviantaba el ansia de claridad universal de Descartes, ocurre en plena vigilia.

CONCLUSIONES

El transhumanismo manifiesta la pretensión de emplazarnos en un contexto en el que, la supeditación de la mediación tecnológica contemporánea, transformada ya en un gigantesco entorno-horizonte de gestión cognoscitiva y pragmática del mundo, trastoca las claves hermenéuticas tradicionales de la subjetividad y de nuestro reconocimiento como ser corporal asido al mundo. Su trasfondo metafísico se aprovecha de una fuerza instrumental expansiva que parece no tener límites o freno alguno para recomponer los marcos de lo viviente (y de todo aquello que es merecedor de preservación) sobre la base de un imperativo de perfección técnica ininterrumpida.

Ahora bien, los férreos lazos de sujeción tecnológica que el transhumanismo reserva para el hombre contemporáneo, no solo supone una tentativa de ofrecer la seguridad de un cobijo primordial frente a la condición de extrañamiento y desamparo radical (*Hilfflosigkeit*) en la que se encuentra, de acuerdo con una de las nociones más certeras e inspiradoras de Freud (2016). El programa ideológico transhumanista cumple a rajatabla las directrices de una previsión escatológica imposible de encubrir. Esta soterrada decantación teleológica se encuentra anticipada ya en Dessauer (1964), uno de los primeros en atreverse a avanzar una reflexión de calado sobre el hecho técnico moderno, cuando compara el avance instrumental con un acto divino trascendente, que puede ser descubierto finalmente por la razón humana gracias a su intrepidez prometeica. En ese sentido, el sometimiento a la

²⁹ Lloyd (2002) plantea que el universo, si se considera un sistema computacional, tendría que haber realizado, desde el *Big Bang*, 10^{120} operaciones en 10^{90} dígitos binarios, algo que excede el número de partículas elementales existente.

absoluta disponibilidad operativa que es consagrada por el proyecto transhumanista responde, en gran medida, a la inconformidad instintiva del hombre actual ante el estado de cosas imperante, a su inclinación por desentenderse de la íntima responsabilidad que le es propia, en última instancia, a ese cansancio alienante del que habla el escritor Handke (2017), en la medida en que disminuye y esclerotiza la vida. De ahí el persistente afán por trascenderla artificialmente, por devaluarla como un rescoldo fútil ante el poder triunfante de la razón bioartefactual (Linares, 2019), por abolirla en toda su integridad a fin de sobreponerse a la vergüenza y a la humillación de nuestro destino precario y fugaz. Visto desde este punto de vista, tal vez no sea posible conjurar del todo las funestas consecuencias derivadas de la mítica caída en la materia (ensomatosis). Quizás no sea posible retornar, como así era el deseo de Bacon (2011), al estado prelapsario erigido por Dios en el jardín del Edén (Anders, 2011), pero sí crear un nuevo ideal antropológico ajustado a las leyes de la biomecánica, un trasunto de paraíso alternativo a medida del hombre.

CONFLICTOS DE INTERÉS

El autor declara que no presenta conflictos de interés financiero, profesional o personal que pueda influir de forma inapropiada en los resultados obtenidos o las interpretaciones propuestas.

REFERENCIAS

- Agatonović, M. (2021). The fiction of simulation: a critique of Bostrom's simulation argument. *AI & Society*. <https://doi.org/10.1007/s00146-021-01312-y>
- Anders, G. (2011). *La Obsolescencia del Hombre. Sobre la Destrucción de la vida en la época de la tercera revolución industrial* (Vol. 2). Pre-textos.
- Bacon, F. (2011). *La Gran Restauración (Novum Organum)*. Tecnos.
- Ball, P. (2012). *Contra natura. Sobre la idea de crear seres humanos*. Turner Publicaciones.
- Bishop, M. P., Elder, S. T., y Heath, R. G. (1963). Intracranial Self-Stimulation in Man. *Science*, 140(3565), 394-396. <https://doi.org/10.1126/science.140.3565.394>
- Bostrom, N. (2003). Are We Living in a Computer Simulation? *The Philosophical Quarterly*, 53(211), 243-255. <https://doi.org/10.1111/1467-9213.00309>
- Bostrom, N. (2017). *Superintelligence: Paths, Dangers, Strategies*. Oxford University Press.

- Chalmers, D. J. (1996). *The Conscious Mind: In Search of a Fundamental Theory*. Oxford University Press.
- Chalmers, D. J. (2005). The Matrix as Metaphysics. En C. Grau (ed.), *Philosophers Explore the Matrix* (pp. 132-176). Oxford University Press.
- Clarke, A. C. (1973). *Profiles of the Future: An Inquiry into the Limits of the Possible*. Harper & Row.
- Delgado, J. M. R. (1969). *Physical Control of the Mind: Toward a Psychocivilized Society*. Harper & Row.
- Descartes, R. (2009). *Meditaciones acerca de la Filosofía Primera. Seguidas de las objeciones y respuestas*. Universidad Nacional de Colombia.
- Descartes, R. (2011). *Discurso del método*. Alianza Editorial.
- Dessauer, F. (1964). *Discusión sobre la técnica*. Rialp.
- Dijksterhuis, E. J. (1969). *The mechanization of the World Picture*. Oxford University Press.
- Dreyfus, H. L. (1992). *What computers Still Can't do: A Critique of Artificial Reason*. The MIT Press.
- Dreyfus, H. L., y Dreyfus, S. E. (1986). *Mind over Machine. The Power of Human Intuition and Expertise in the Era of the Computer*. Free Press.
- Du Bois-Reymond, E. (1872). *Über die Grenzen des Naturerkennens. Ein Vortrag in der zweiten öffentlichen Sitzung der 45. Versammlung deutscher Naturforscher und Ärzte zu Leipzig am 14. August 1872*. Veit.
- Feigl, H. (1958). The "Mental" and the "Physical". En H. Feigl, M. Scriven, G. Maxwell (eds.), *Concepts, Theories, and the Mind-Body Problem* (Vol. 2, pp. 370-497). University of Minnesota Press. <https://hdl.handle.net/11299/184614>
- Finkelstein, G. (2013). *Emil du Bois-Reymond. Neuroscience, Self, and Society in Nineteenth-Century Germany*. The MIT Press.
- Floridi, L. (2009). Against digital ontology. *Synthese*, 168(1), 151-178. <https://doi.org/10.1007/s11229-008-9334>
- Freud, S. (2016). *Inhibición, síntoma y angustia*. Amorrortu editores.
- García Selgas, F. J. (1994). El «cuerpo» como base del sentido de la acción. *Reis: Revista Española de Investigaciones Sociológicas*, (68), 41-83. <https://doi.org/10.2307/40183757>

- Gödel, K. (1931). Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme I. *Monatshfte für Mathematik und Physik*, 38, 173-198. <https://doi.org/10.1007/BF01700692>
- Han, B.-C. (2014). *Psicopolítica*. Herder Editorial.
- Handke, P. (2017). *Ensayo sobre el cansancio*. Alianza Editorial.
- Hochreiter, S., y Schmidhuber, J. (1997). Long Short-Term Memory. *Neural Computation*, 9(8), 1735-1780. <https://doi.org/10.1162/neco.1997.9.8.1735>
- Huxley, A. (1932). *Brave New World*. Chatto & Windus.
- Jung, C. G., y Kerényi, K. (2004). *Introducción a la esencia de la mitología*. Ediciones Siruela.
- Kim, J. (2000). *Mind in a Physical World. An Essay on the Mind-Body Problem and Mental Causation*. The MIT Press.
- Kim, J. (2005). *Physicalism, or Something Near Enough*. Princeton University Press.
- Kurzweil, R. (2005). *The Singularity Is Near: When Humans Transcend Biology*. Penguin Group.
- Larson, E. J. (2021). *The Myth of Artificial Intelligence. Why Computers Can't Think the Way We Do*. The Belknap Press of Harvard University Press.
- Levesque, H. J. (2014). On Our Best Behavior. *Artificial Intelligence*, 212, 27-35. <https://doi.org/10.1016/j.artint.2014.03.007>
- Linares, J. E. (2019). *Adiós a la naturaleza. La revolución bioartefactual*. Plaza y Valdés.
- Lloyd, S. (2002). Computational Capacity of the Universe. *Physical Review Letters*, 88(23), 237901. <https://doi.org/10.1103/PhysRevLett.88.237901>
- Lloyd, S. (2006). *Programming the Universe: A Quantum Computer Scientist Takes On the Cosmos*. Vintage Books.
- Meyrink, G. (2010). *El Golem*. Plutón ediciones. (Trabajo original publicado en 1915).
- Morozov, E. (2015). *La locura del solucionismo tecnológico*. Clave Intelectual.
- Mumford, L. (2016). *El pentágono del poder. El mito de la máquina* (Vol. 2). Pepitas de Calabaza.
- Neumann, J. V. (1966). *Theory of Self-Reproducing Automata*. University of Illinois Press.

- Nöe, A. (2009). *Out of Our Heads. Why You Are Not Your Brain, and Other Lessons from the Biology of Consciousness*. Hill and Wang.
- Nordmann, A. (2012). Object lessons: towards an epistemology of technoscience. *Scientiae Studia*, 10(especial), 11-31. <https://doi.org/10.1590/S1678-31662012000500002>
- Olds J., y Milner, P. (1954). Positive reinforcement produced by electrical stimulation of septal area and other regions of rat brain. *Journal of Comparative and Physiological Psychology*, 47(6), 419-427. <https://psycnet.apa.org/doi/10.1037/h0058775>
- Peirce, C. S. (1929). Guessing. *The Hound and Horn*, 2(3), 267-282.
- Putnam, H. (1988). *Razón, Verdad e Historia*. Tecnos.
- Reichenbach, B. R. (1978). Monism and the Possibility of Life after Death. *Religious Studies*, 14(1), 27-34. <https://doi.org/10.1017/S0034412500010465>
- Ryle, G. (2005). *El concepto de lo mental*. Ediciones Paidós Ibérica.
- Sierra, C. H. (2021). Medical imaginaries and the future horizons of consciousness physiology in the work of Stanisław Lem. *Rassegna Italiana di argomenti polacchi*, (12), 39-53. <https://plitonline.it/2021/plit-12-2021-39-53-carlos-hugo-sierra>
- Sloterdijk, P. (2000). *Normas para el parque humano. Una respuesta a la Carta sobre el humanismo de Heidegger*. Ediciones Siruela.
- Sloterdijk, P. (2012). *Haz de cambiar tu vida. Sobre antropotécnica*. Pre-Textos.
- Smart, J. J. C. (1959). Sensation and Brain Processes. *The Philosophical Review*, 68(2), 141-156. <https://doi.org/10.2307/2182164>
- Szocik, K., Tkacz, B., y Gulczyński, P. (2020). The revelation of superintelligence. *AI & Society*, 35(3), 755-758. <https://doi.org/10.1007/s00146-020-00947-7>
- Steinhart, E. (2010). Theological Implications of the Simulation Argument. *Ars Disputandi*, 10(1), 23-37. <https://doi.org/10.1080/15665399.2010.10820012>
- Tallis, R. (2008). *The Kingdom of Infinite Space. A Portrait of Your Head*. Yale University Press.
- Tallis, R. (2016). *Aping Mankind. Neuromania, Darwinitis and the misrepresentation of humanity*. Routledge.

Turing, A. M. (1950). Computing Machinery and Intelligence. *Mind*, 59(236), 433-460. <https://doi.org/10.1093/mind/LIX.236.433>

Vinge, V. (1993). The Coming Technological Singularity: How to Survive in the Post-Human Era. *Vision 21: Interdisciplinary Science and Engineering in the Era of Cyberspace*, 11-22. <https://ntrs.nasa.gov/citations/19940022856>

Vogt, C. (1874). *Physiologische Briefe für Gebildete aller Stände*. Ricker.

Wells, H. G. (1896). *The Island of Doctor Moreau*. Heinemann.

Zamiatin, E. (1924). *We*. Dutton.

Zuse, K. (1969). *Rechnender Raum*. Friedrich Vieweg & Sohn.